A Transgenic Model for Purification of Higher-Order RNA Transcription Complexes

Dissertation
zum Erlangen des Doktorgrades
der Naturwissenschaften

Vorgelegt beim Fachbereich Biochemie, Pharmazie,
und Lebensmittelchemie
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von
Erick Berglund
aus Deerfield, New Hampshire (USA)

Frankfurt am Main (1999)
(DF1)

vom Fachbereich Biochemie, Pharmazie, und Lebensmittelchemie der
Johann Wolfgang Goethe-Universität als Dissertation angenommen

Dekan:	Prof. Dr. A. Mosandl
Erster Gutachter:	Prof. Dr. Dr. H. Fasold
Zweiter Gutachter:	Prof. Dr. B. Ludwig
Datum der Disputation:	14. Juli, 1999

vom Fachbereich Biochemie, Pharmazie, und Lebensmittelchemie der
Johann Wolfgang Goethe-Universität als Dissertation angenommen

Dekan:	Prof. Dr. A. Mosandl
Erster Gutachter:	Prof. Dr. Dr. H. Fasold
Zweiter Gutachter:	Prof. Dr. B. Ludwig
Datum der Disputation:	14. Juli, 1999

TABLE OF CONTENTS

TABLE OF CONTENTS

1. ZUSAMMENFASSUNG

Die Expression jedes Protein kodierenden Gens im Genom von Eukaryonten ist ein dynamischer, individuell regulierter Prozeß. Die ihm zugrundeliegenden Kontrollfunktionen sind für die Biologie der Eukaryonten von entscheidender Bedeutung, sind von ihnen doch die spezifische Differenzierung und Entwicklung von Geweben abhängig. Zusätzlich fungiert dieses Kontrollsystem als Endpunkt diverser Signalübertragungswege und steuert die Umgestaltung der Zelle und ihrer Funktionen als Reaktion auf wechselnde metabolische Bedürfnisse und äußere Einflüsse. Der Fluß der genetischen Information von DNA zu Boten-RNA zum Protein unterliegt einer Vielzahl regulatorischer Prozesse, doch spielt bei der Genexpression die Einleitung der Transkription die zentrale Rolle. Die Initiation der Transkription findet dann statt, wenn sich der RNA Polymerase II (RNA Pol II) Komplex in korrekter Orientierung am Anfang des Bereichs eines Gens in der Erbsubstanz anlagert. Von dort gleitet sie an den Anfang der kodierenden Region des Gens und bewegt sich dann an der DNA entlang, wobei sie eine entsprechende RNA-Kopie erstellt.

Die individuelle Natur der genetischen Regulation wird durch bestimmte Abschnitte in der Erbsubstanz reguliert, den Promotoren. Diese sind generell zweigeteilt, bestehend aus einem Kernelement und für jedes Gen einzigartigen regulatorischen Bereichen. Die Kernregion ist in den meisten eukaryontischen Promotoren konserviert und dient der korrekten Orientierung der RNA Pol II. Sie liegt in der Regel in der Nähe des kodierenden Abschnitts und besteht meist aus einem Bereich den man auch als „TATA-Box" bezeichnet und der Startstelle für die Transkription. Von diesem sogenannten Kernpromotor aus beginnt die RNA Pol II an der DNA entlangzuwandern. Im Gegensatz dazu sind die regulatorischen Bereiche extrem variabel und spezifisch für jedes individuelle Gen. Diese Regionen enthalten ganz bestimmte Kombinationen kurzer DNA Sequenzen über welche die Transkription entweder stimuliert (Verstärker oder „Enhancer") oder gehemmt (Drosselelemente oder „Silencer") werden kann. Diese Elemente können weit vom Kernpromotor entfernt liegen, teilweise in einem Abstand von vielen tausend Basenpaaren vom Kernpromoter entfernt. Die an solche Promotorelemente bindenden regulatorischen Proteine bezeichnet man als Transkriptionsfaktoren.

Neuere Arbeiten auf dem Gebiet der RNA Transkription entdeckten eine Vielzahl sogenannter „Genereller Transkriptionsfaktoren" (GTF), die für die akkurate Rekrutierung der RNA Pol II am Kernpromotor unentbehrlich sind. Diese Rekrutierung wird durch die Anlagerung des generellen Transkriptionsfaktors TFIID an die TATA-Box initiiert. Nach der Bindung von TFIID, lagern sich dort auch andere GTF in einer festgelegten Reihenfolge an. Hierbei interagieren sie sowohl miteinander als auch mit der DNA und bilden den sogenannten „Präinitiationskomplex" („Preinitiation Complex", PIC). Nur wenn dieser Komplex erfolgreich gebildet wird, kann die Polymerase die Transkription des Gens beginnen. Er ist unerläßlich für die Transkription eukaryontischer Gene damit die RNA Pol II eine

geringe basale Transkriptionsrate erreicht. Die genaue, schrittweise Vollendung des PIC wird durch regulatorische Proteine beeinflußt die an die Verstärker- bzw. Drosselelemente des Promotors binden. Solche Aktivatoren oder Repressoren beeinflussen das Zustandekommen des PIC durch direkte oder indirekte Protein-Protein Wechselwirkungen. Somit gibt es zwei Hauptklassen von Transkriptionsfaktoren, die Basalfaktoren zur Einleitung der Transkription und andere Proteine, Aktivatoren und Repressoren, welche die Geschwindigkeit bestimmen, mit der der Basalkomplex die Transkription in Gang setzt. Der einzige GTF mit der Fähigkeit selbst spezifische Sequenzen innerhalb des Kernpromotors zu erkennen ist TFIID.

TFIID ist ein Multi-Proteinkomplex bestehend aus einem Kernprotein, dem „TATA-Binding- Protein" (TBP) und mindestens zwölf verschiedene „TBP-Associated-Factors" ($TAF_{II}s$). Der gesamte TFIID Komplex spielt in der Regulation der Genexpression, also im Fluß der genetischen Information von DNA zu Boten-RNA eine entscheidende Rolle, da gezeigt werden konnte, daß er für die regulierte Transkription *in vitro* notwendig ist. Zusätzlich sind auch einzelne $TAF_{II}s$ in eine Reihe regulatorischer Funktionen mit Einfluß auf die Transkription involviert. Diese beinhalten Protein-Protein Wechselwirkungen mit an Verstärkerelemente gebundenen Aktivatoren, PIC bildende und stabilisierende Funktionen, Histon Acetylierung und die Sensitivität für Phosphorylierungen durch wichtige Zellzyklus abhängige Kinasen. Es besteht Grund zur Annahme, daß die Hauptfunktion des TFIID Komplexes in der Integration einer Vielzahl positiver und negativer regulatorischer Signale für die Formation des PIC und die Einleitung der Transkription besteht. Interessanterweise wurden in jüngster Zeit verschiedene TFIID Komplexe beschrieben, die aus unterschiedlichen Kombinationen von $TAF_{II}s$ und TBP zusammengesetzt sind. Weiterführende Untersuchungen an Zellinien aus unterschiedlichen Geweben legen die Vermutung nahe, daß manche $TAF_{II}s$ gewebespezifisch exprimiert werden. Dies läßt wiederum den Schluß zu, daß solche gewebespezifischen $TAF_{II}s$ eine Schlüsselrolle bei der auf bestimmte Gewebe beschränkten Expression von Genen zukommt. Gleichwohl sind Daten die dieses Model unterstützen rar, was eventuell mit der geringen Auswahl an Geweben zu tun hat, die in Form von immortalisierten Zellinien zur Verfügung stehen.

Mit der vorliegenden Arbeit wurde die Basis für die Umgehung dieser Probleme geschaffen, nämlich die Möglichkeit der simultanen Untersuchung von TFIID aus allen Geweben der Vertebraten. Sie beschreibt die Herstellung einer Linie transgener Mäuse, welche die cDNA für menschliches TBP tragen. Zusätzlich wurde die cDNA mit zwei Peptidmotiven gekoppelt, welche die affinitätschromatographische Reinigung des transgenen TBP über zwei unterschiedliche Epitope an dessen N-Terminus erlauben. Das Erste ist das neun Aminosäuren lange Hämagglutinin (HA) Epitop für die immuno-affinitätschromatographische Reinigung mit einem kommerziell erhältlichen monoklonalen anti-HA Antikörper. Der zweite Marker besteht aus sechs aufeinanderfolgenden Histidinen, welche Komplexe mit positiv geladenen Schwermetallionen bilden und routinemäßig über Ni^{2+}-Agarosesäulen

gereinigt werden können. Da der generelle Transkriptionsfaktor (GTF) TFIID einen Komplex bestehend aus TBP und verschiedenen $TAF_{II}s$ darstellt, ermöglicht die Markierung des transgenen TBP mit den beiden Epitopen eine neue effektive Strategie die individuellen Komponenten des TFIID Komplexes in Geweben besser zu charakterisieren.

Mauslinien die das markierte transgene TBP exprimierten wurden für die Präparation von Kernextrakten aus Leber, Gehirn und Niere benutzt. Zunächst wurde TFIID aus diesen Kernextrakten durch Ionenaustauschchromatographie angereichert. Danach wurde der TFIID Komplex durch Nickel-Affinitätschromatographie über das Histidine Epitop am transgenen TBP weiter angereichert. Diese partiell gereinigte TFIID Fraktion zeigte bereits spezifische Aktivität in *in vitro* Transkriptionsversuchen. Dies zeigt einerseits, daß das transgene TBP in der Tat in seiner aktiven Form in den untersuchten transgenen Geweben vorliegt. Weiterhin ist es in der Lage, mit endogenen Maus-$TAF_{II}s$ einen funktionalen TFIID Komplex zu formen. Die Ergebnisse dieser Arbeit lassen folglich den Schluß zu, daß die eingeführten Epitope zur Markierung des transgenen TBP zur effizienten Reinigung von TFIID in einer transkriptionsaktiven Form aus unterschiedlichen Geweben geeignet ist.

Obwohl in den letzten Jahren diverse neue, effektive Strategien zur Reinigung von TFIID aus Zellkulturen entwickelt wurden, um dessen individuelle Komponenten besser zu charakterisieren, konnte in dieser Arbeit zum ersten mal eine solche Reinigungsmethode auf mehrzellige Organismen übertragen werden. Der offensichtliche Vorteil dieses Ansatzes ist es, dem Forscher den direkten Zugang zu gewebespezifischen TFIID Komplexen und deren TAF_{II} Untereinheiten zu ermöglichen. Dieses affinitätsgereingte TFIID kann nun für die qualitative und mechanistische Untersuchung von gewebe-, entwicklungs- oder umgebungsspezifischer genetischer Regulation in mehrzelligen Organismen herangezogen werden.

Mit diesem neu erworbenen Wissen könnte man zusätzlich neue Strategien zur Bekämpfung lebensbedrohlicher Krankheiten entwickeln, bei denen eine übermäßige oder eine unzureichende Transkription eines Gens die Ursache ist. Die angestrebte Identifikation neuer gewebespezifischer Transkriptionsfaktoren aus der in dieser Arbeit beschriebenen transgenen Maus, eröffnet möglicherweise den Weg zur therapeutischen Intervention gewebespezifischer Genregulation und könnte eine Plattform für die Entwicklung spezifischerer Medikamente sein.

2. BACKGROUND

2.1 Introduction

Eukaryotic organisms have developed exquisitely refined systems to modulate genetic expression and to precisely execute the vast array of genetic programs which guide complex biological events. Unique events such as the steps of development and differentiation must be accomplished successively in a well-controlled temporal and spatial fashion. Ongoing processes such as the day to day reshaping of the biochemical profile of the cell must be in measured response to a wide spectrum of environmental conditions and challenges. The current general view is that the first phase of genetic expression, the transcription of DNA to RNA, represents a key step at which expression is predominantly regulated. Therefore, there has been considerable effort over the last 15 years to dissect and deduce the molecular mechanisms which govern and guide this essential event.

2.2 A three polymerase system

A major evolutionary distinction between prokaryotes and eukaryotes is the transition from a single enzyme for transcribing DNA into RNA (RNA Polymerase) to three separate enzymes. The three multimeric eukaryotic RNA polymerases (RNA Pol I, II, and III) are all related in that they retain sequence homology in the largest subunit (1), and actually have five subunits in common (2)-(3;4). Although the RNA Pol enzymes are complex and very large multi-subunit factors (the complete RNA Pol II complex is ~500 kDa) they contain no intrinsic ability to select appropriate transcriptional start sites within DNA. This is illustrated well by the Pol II enzyme, which can efficiently bind DNA *in vitro* in a non-specific manner and initiate RNA production from random start sites on a DNA template (5). Therefore, it is not surprising that all three eukaryotic RNA polymerase enzymes require individual and complex accessory mechanisms to recognize their respective transcription initiation sites. This is in marked contrast to prokaryotes such as *E. coli*, which require only a single polypeptide (sigma factor) to interact with its sole RNA Pol complex and faithfully recruit it to its correct transcription start sites in the genome (6). It is thought that eukaryotes have developed a more complex and modulated system over that of prokaryotes owing to their need to accurately regulate transcription within a more complicated genome. There is a distinct division of labor among the three eukaryotic transcription systems, which somewhat corresponds to the three major kinds of RNA species present. It is exclusively RNA Pol II that plays the central role of transcribing the large number of individual protein coding genes into messenger RNA (mRNA). The comparatively small subset of non-protein coding genes, mainly translational RNA and ribosomal RNA (tRNA and rRNA), are produced by the combined activities of RNA Pol I and RNA Pol III. While this may seem a narrow task, it should be noted that as much as 4/5 of the total RNA being synthesized in a rapidly growing cell is rRNA. Nevertheless, it is understandable that the overwhelming majority of efforts to understand genetic regulation have been focussed on protein-coding genes, which are exclusively recognized by RNA Pol II.

2.3 Promoter structure

The advent of molecular biology brought early recognition that some portions of a given gene's non-coding sequence (generally the 5' end) are indispensable for its correct regulation and transcriptional initiation. These "control regions" are often referred to in a general sense as promoters and presumably contain all of the information related to the correct expression of the associated gene. Promoters can be separated into at least two functional categories: a core promoter and upstream (or downstream) regulatory elements. The core promoter specifies the accurate initiation site for transcription, whereas the regulatory elements, often referred to as enhancers, direct gene activation and repression in response to spatial, temporal and environmental influences.

Virtually all RNA Pol II core promoters (also referred to as "type II" promoters) contain one or both of two common sequence features which can function independently or synergistically. The vast majority of type II promoters characterized so far contain a so-called "TATA Box" sequence (consensus TATAa/tAa/t), which is normally located about 25 base pairs upstream from the start site in higher eukaryotes (7). In addition, many promoters have a pyrimidine-rich initiator "Inr" sequence (consensus YYANt/aYY) which is located very near the transcription start site (reviewed in 8). These two sequence elements are thought to serve as tethering points for the basic transcriptional machinery, a stereo-specific complex of "general transcription factors" (GTFs) and the RNA Pol II enzyme. This protein structure in its entirety (GTFs and polymerase) is often referred to as the "pre-initiation complex" (PIC), and is able to initiate accurate transcription of RNA when fully assembled upon the core promoter.

The regulatory elements of promoters are much more diverse in their architecture and occurrence. There have to date been described a great many such elements, a unique combination and arrangement of which are associated with each individual gene. They may appear within or near the core promoter (promoter proximal elements), as well as far up- or downstream of the start site (enhancers). These regulatory elements serve as binding sites for gene-specific nuclear factors, which in turn modulate the level of transcriptional initiation. It is presumed that, in large part, control over any specific gene is effected through the presence (or absence) of these tethered factors, which influence the rate of assembly and transcriptional activity of the PIC on the core promoter.

2.4 Basal and activated transcription *in vitro*: different elements of an *in vivo* mechanism

Since the demonstration of multiple RNA Pol II accessory factors (GTFs) almost 20 years ago (9), a flurry of biochemical and genetic studies have revealed a detailed picture of the events that govern transcriptional initiation at the core promoter. Six minimal factors for accurate Pol II activity, TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH were identified, originally as chromatographic fractions. Since then, in most cases, the cDNAs for all the subunits have been identified. These factors are evolutionarily conserved in diverse species, including yeast and human. When all six factors are combined with RNA Pol II and relaxed core promoter DNA, it is possible to generate low levels of accurately initiated RNA *in vitro*, which is referred to as "basal" transcription (see figure 2.1, top). Such minimal transcription

can occur only after the individual factors assemble a stereo-specific complex, apparently via a stepwise pathway (discussed below). During this assembly, RNA Pol II is incorporated onto the promoter via interactions with the forming complex. RNA transcription can commence thereafter, upon completion of the PIC. It should be noted that natural constraints, such as chromatin packaging and repressor factors, are presumed to hinder the assembly function of these six GTFs factors *in vivo*. An isolated core promoter would not, in all likelihood, be able to function as such by itself *in vivo*. Nevertheless, the ability to recreate basal RNA Pol II transcription from core promoters *in vitro* has made it possible to define the minimal aspects of this process, which is a necessary prerequisite to understanding the superimposed regulatory systems *in vivo*.

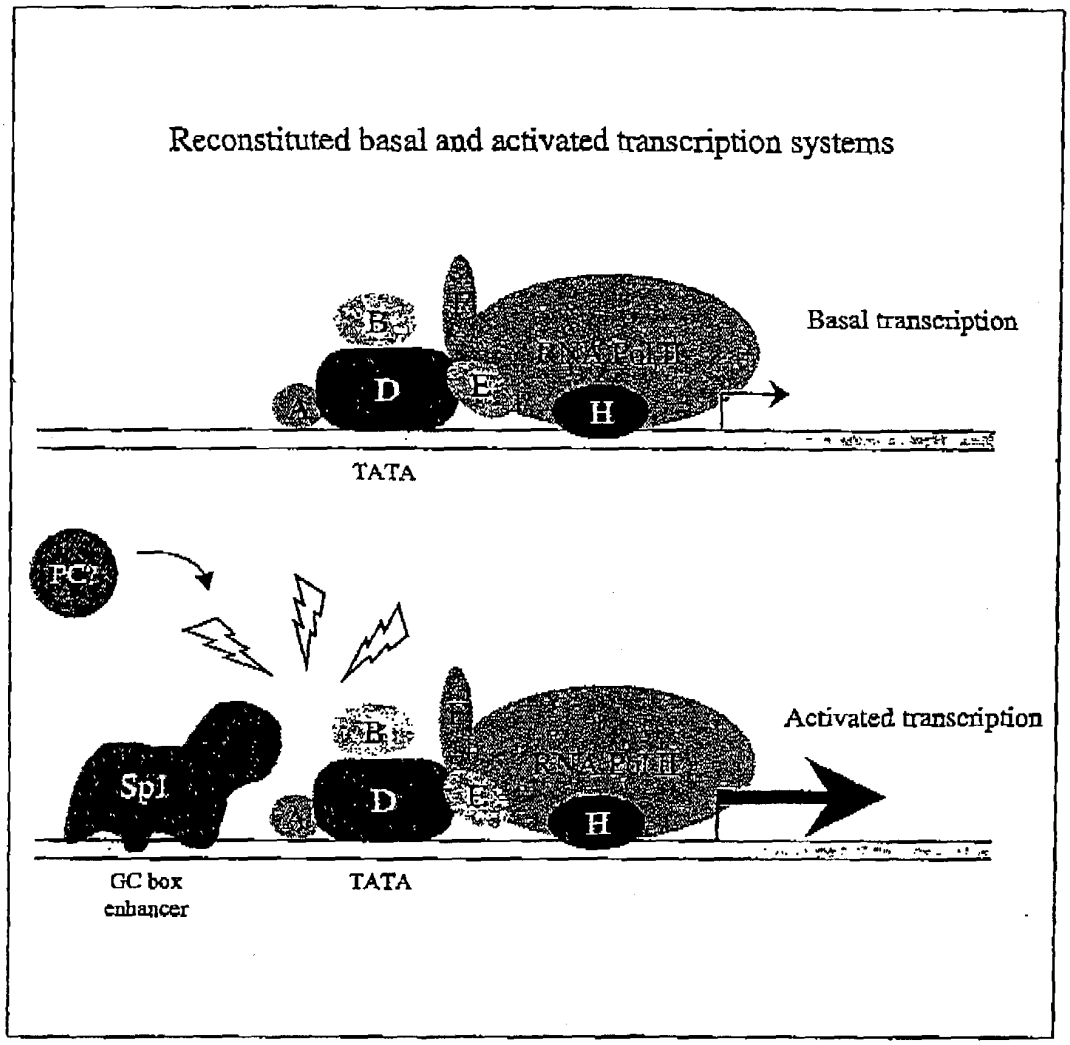


Figure 2.1: Schematic comparison of the elements which comprise basal and activated *in vitro* transcription. Arrows indicate transcription start site. Size of arrow indicates relative level of RNA transcript produced. The GC box is a typical upstream enhancer element. Reconstitution of some systems from highly purified components (e.g. human), can require the presence of one or more "positive cofactors"(PC) for enhancer-directed activation.

By comparison, *in vitro* experimental systems using core promoters combined with regulatory elements are also possible. The minimal PIC components may be combined with positive regulatory factors (or a chromatographic fraction containing such) to recreate "activated" transcription from a promoter containing both a core region and one or more enhancer elements (see figure 2.1, bottom). With these further ingredients, activated *in vitro* transcription can produce RNA anywhere from 2 to 100 fold more efficiently than a corresponding basal reaction. Transcriptional activation studies in some systems (e.g. human) often require the inclusion of one or more "positive co-factors" (PCs) to capacitate an activation effect when using highly purified GTF components (reviewed in "10). This class of factors

is clearly distinguished from enhancer-bound activators. The principles of activated transcription is currently the topic of intense study, as this phenomenon is assumed to be a pivotal aspect of genetic regulation as it occurs *in vivo*. Considerable evidence strongly suggests that at least part of this mechanism involves DNA-bound regulatory factors interacting with components of the PIC. It is proposed that such interactions can recruit GTFs to the core promoter, thus accelerating an otherwise slower, rate-limiting step (discussed below).

2.5 The basal machinery

It is unlikely that the core promoters for any given RNA Pol II-transcribed genes within a single organism are alike. Available sequence data suggests they are as many and varied as there are genes in the genome. Two well-defined core promoter consensus sequences are mentioned above, the TATA box and the Inr element. Both of these elements are present in potent forms on the major late promoter of adenovirus (AdML), upon which much of the study of human PIC formation was initiated (11). The clearest picture of the fundamentals of human PIC assembly comes from studies using the minimal set of purified factors and the AdML promoter, but it is worthwhile to bear in mind that this picture is only one of a spectrum of variations which occur generally on core promoters. Naturally occurring promoters can have weak or deleted versions of either of these elements, yet still retain transcriptional competency. Parallel assembly mechanisms for TATA-less promoters have been proposed, but are not nearly as well understood as those for TATA-containing promoters (reviewed in 12).

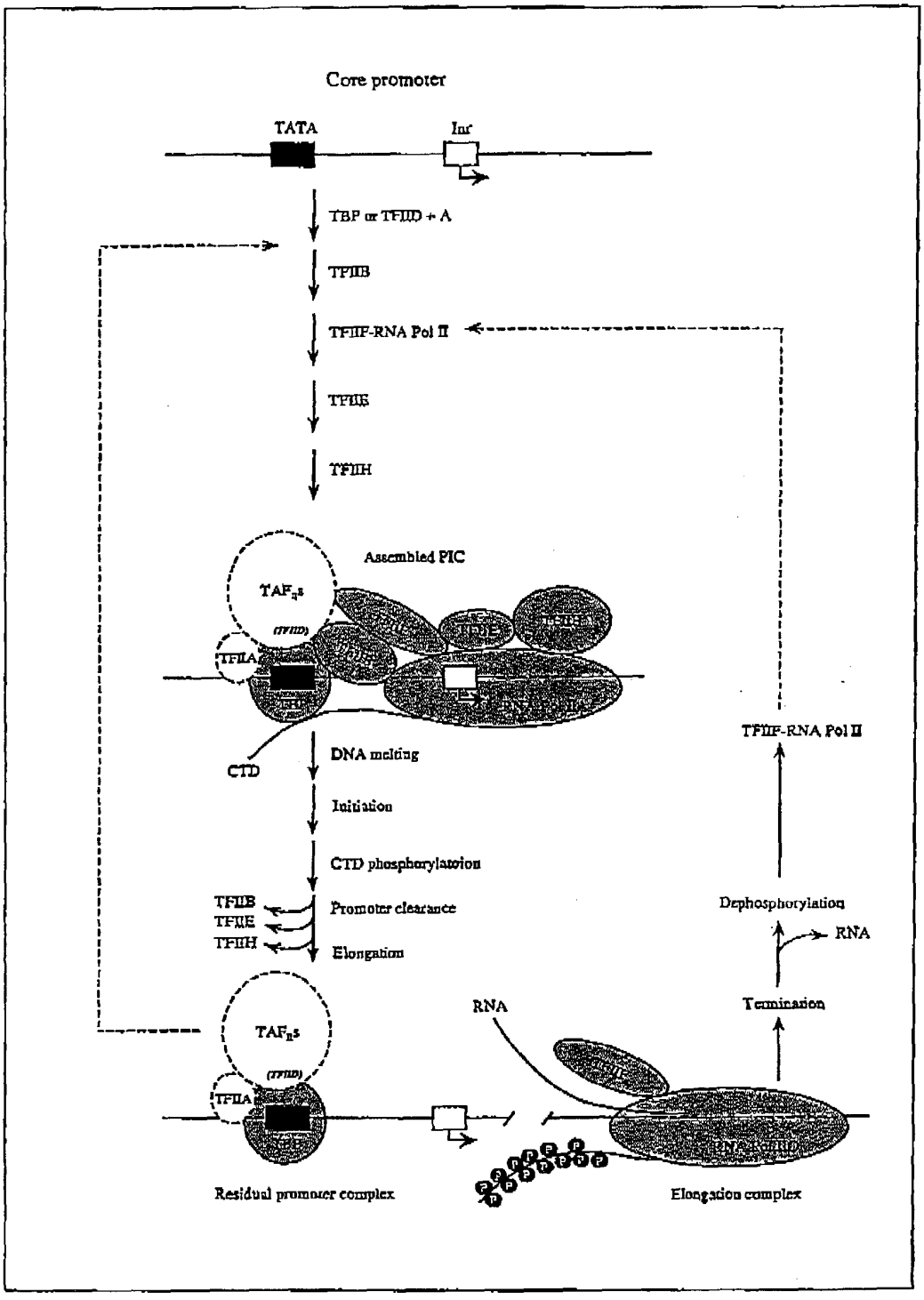


Figure 2.2: Schematic diagram of stepwise PIC assembly on a TATA-containing core promoter. Grey colored GTFs are considered essential for basal activity. Round black solids indicate phosphorylation. (13).

In the most general case, PIC formation begins when the factor TFIID recognizes and binds to the TATA-box sequence in the core promoter (see top of figure 2.2). TFIID is the only GTF that has sequence specificity, and its preliminary binding appears to be a critical rate-limiting step of PIC formation (14-17) . This sequence recognition is provided exclusively by one of its subunits, called TATA-binding protein (TBP). Currently there are at least a dozen other known TFIID subunits which can associate through contact with TBP. These TFIID subunits are generally referred to as "TBP-associated factors" or TAF_{II}s. As with the other components of the PIC, strong homologies exist both in TBP and TAF genes from yeast, human and *Drosophila* (see figures 2.3 and 2.4). Although TBP alone can support basal levels of *in vitro* transcription from strong core promoters (such as AdML), the holo-complex of TFIID is generally considered to be one of the basal factors because many weaker promoters require the presence of some TAF_{II}s (discussed below). It is through contacts between TBP and the minor groove of the DNA that TFIID binds to the promoter, causing a severe, anti-helical distortion and bringing about a dramatic bend in the DNA (18-20). It is thought that this event facilitates regulatory mechanisms by bringing upstream and downstream sequences into proximity with one other (reviewed in "13).

TFIIA is normally the next factor to join the assembly, binding stably to the TFIID-DNA complex. Human TFIIA is composed of three subunits of 12, 19, and 35 kDa (reviewed in 21). The holo-TFIIA contacts the TBP molecule as well as the adjacent DNA sequence immediately upstream of the TFIID complex. As a GTF, TFIIA appears to function primarily as a stabilization factor, enhancing the contacts between TFIID and the promoter via contacts both with TBP and certain TAF_{II}s. Like the TAF_{II}s, however, TFIIA is not absolutely required for basal transcription on potent TATA-containing core promoters. However, it is generally included in discussions about the basal GTFs since there are indications that weaker core promoters require TFIIA to stabilize otherwise unfavorable TFIID binding (13;22).

The 35kDa TFIIB monomer is the next factor to join the complex, recognizing and binding to the TFIID-DNA complex. Similar to TFIIA, TFIIB is also able to stabilize weak TBP-TATA interactions (23;24), but can enter the PIC regardless of the presence of TFIIA. TFIIB-DNA contacts occur both upstream and downstream of the TATA box, but apparently not with TFIIA (13). Although TFIIB and TFIIA contact common upstream base pairs, the contacts are on opposite sides of the DNA strand (13). Perhaps the most critical aspects of TFIIB presence in the PIC are its dual roles of recruiting the RNA Pol II-TFIIF complex (see below) to the core promoter through direct contact (25) and positioning of the polymerase for accurate start site selection (reviewed in "26)(27)(28).

Human TFIIF is a heterodimer with subunits of masses 30 and 74 kDa (29). Among the GTFs, TFIIF is unique in its ability to form a stable complex with RNA Pol II (30) It is through this connection that accurate recruitment of the polymerase holoenzyme to the start site occurs, mediated by TFIIB-TFIIF as well as TFIIB-RNA Pol II interactions (26;27). Both of the TFIIF subunits (RAP30 and

RAP74) have additionally been shown to play a more indirect role by reducing RNA Pol II binding to non-specific sites on the DNA.(31;32).

TFIIE is an $\alpha_2\beta_2$ heterotetramer with subunits of 34 and 57 kDa (reviewed in "33) which joins the PIC through direct contacts with Pol II and possibly with TFIIF and TFIID (27;32;34). As with other GTFs, TFIIE assists in further stabilization of the PIC. TFIIE is further involved with regulating the helicase activity of TFIIH (see below) and possibly with melting the DNA to expose a template for transcription initiation (35).

Binding of the large TFIIH through direct contacts with TFIIE (34;36) completes the assembly of the PIC. TFIIH is a complex of nine subunits, ranging in size from 39 to 89 kDa (reviewed in "37). At least three catalytic activities have been attributed to TFIIH, including an ATPase, a bi-directional DNA helicase, and a serine/threonine kinase (reviewed in "38). In addition, at least two TFIIH subunits (ERCC2 and ERCC3) are also components of the DNA excision repair machinery, apparently implicating TFIIH in DNA repair (reviewed in "39). Although differences of opinion exist about the relevance of DNA repair to RNA transcription, it is clear that the phosphokinase and 3'-5' helicase activities are vital to transcript initiation. It is thought that the energy-requiring step of opening the DNA for access of the RNA Pol II is provided by these TFIIH functions (40). In addition, the kinase activity of TFIIH plays a further role by phosphorylating the "C-terminal domain" (CTD) of the largest of the RNA Pol II subunits. The CTD is a highly conserved repeat domain that becomes hyperphosphorylated during the transition from PIC formation to RNA production (elongation) (41).

When the PIC assembly is complete (middle of figure 2.2), separation of the two DNA strands occurs at the start site, the CTD of RNA Pol II becomes phosphorylated, and transcription of RNA from the exposed antisense template commences. Subsequent RNA elongation obligates the polymerase to move along the DNA away from the PIC (promoter clearance). As it moves away only TFIID and TFIIA are left in their positions on the core promoter, still available to support accretion of another PIC (bottom of figure 2.2) (42). TFIIB and TFIIE have been observed to be the first GTFs to dissociate from the PIC, followed soon after by TFIIH (43;44). TFIIF remains associated with the polymerase holoenzyme through RNA production and, ultimately, termination (42).

As mentioned above, the steps toward PIC formation outlined here are an amalgam of numerous biochemical and structural studies on a range of transcription systems, most notably that of human factors assembling on the AdML core promoter. It is important to note that this is not necessarily the only pathway by which a PIC can be assembled. Recently, large multi-protein complexes containing RNA Pol II and most of the GTFs (excluding TFIID and TFIIB) have been purified from nuclear extracts from yeast and mammalians (45;46) It may be that such "RNA Pol II holoenzymes" may be alternatively used *in vivo* for a more rapid PIC assembly pathway by binding in a single step to the promoter-bound TFIID-TFIIB complex, resulting in a higher rate of initiation (47). Hence, a minimal pathway as described here should be viewed as a flexible framework for PIC assembly, with individualized variations occurring on other promoters and under different conditions.

2.6 TFIID - a key factor

Being the initial factor to bind the core promoter, TFIID was widely recognized to play a crucial role in transcription soon after it was initially identified as a chromatographic fraction from human nuclei (9). Early studies of activated *in vitro* transcription also identified TFIID as a target for gene-specific activators and enhancers (11;48;49), further underscoring its apparent pivotal role in regulated transcription. However, subsequent efforts to isolate a single human peptide with TATA binding ability long proved fruitless. A breakthrough came when several groups simultaneously purified and cloned a 25 kDa yeast protein with a specific TATA-box affinity (TBP).(50-53). Thereafter, the yeast TBP sequence was rapidly used to identify and clone homologous TBP genes from *Drosophila* (54;55) , human (56-58), mouse (59), and a variety of other higher eukaryotes (reviewed in "60).

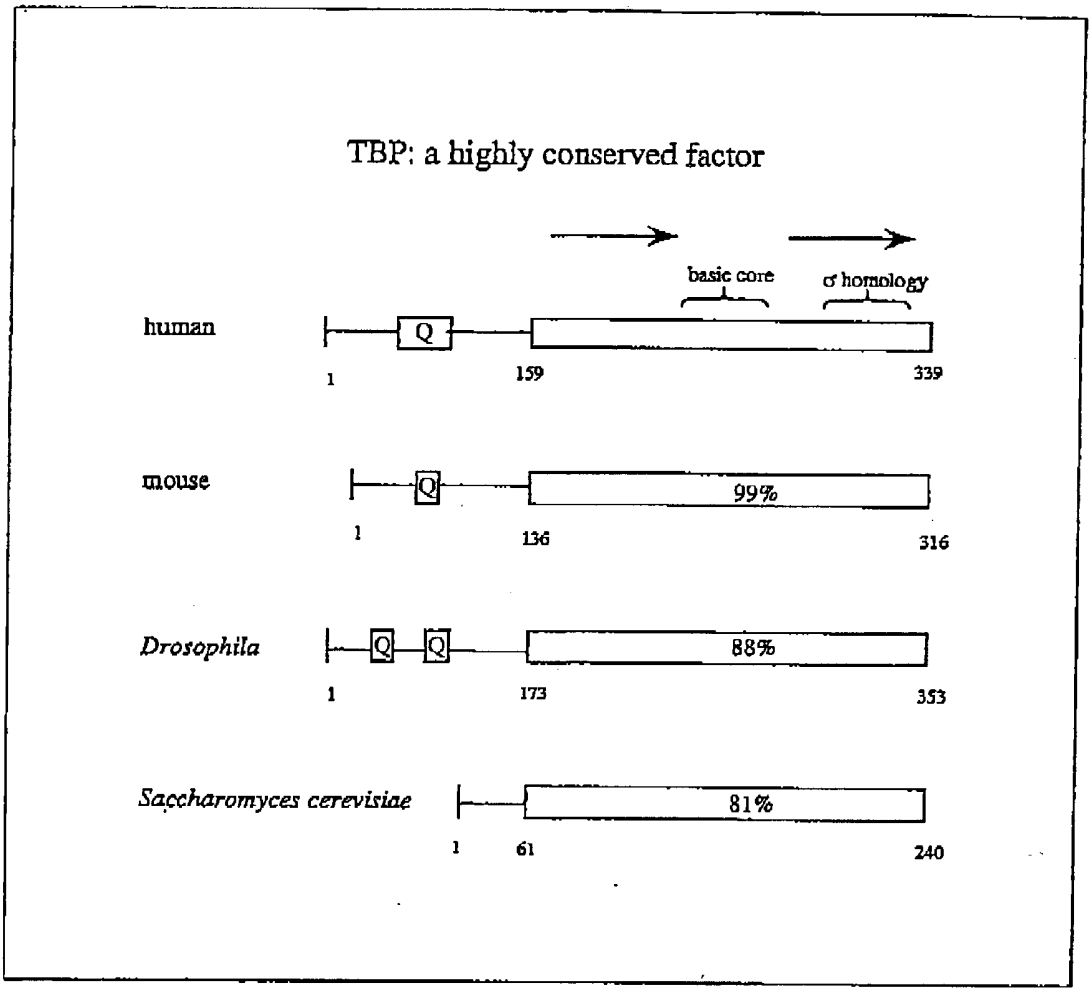


Figure 2.3: Comparison of TBP primary structure from different species. Glutamine stretches are marked with "Q". Conserved basic core and homology to bacterial σ factor indicated with brackets. Arrows indicate direct repeats. Sequence homology to human carboxy-terminal domain represented as a percentage.

Comparison of the TBP sequences from various species reveals a protein that can be divided into two domains: a highly conserved C-terminal domain and a variable N-terminal domain (see figure 2.3). Among some of the conserved features of the C-terminal domain include two interrupted direct repeats of some 80 amino acids and a central core rich in positively charged lysine residues (reviewed in 61) (reviewed in 62). There is also a region with sequence similarity to the promoter-binding bacterial σ factor. Binding studies indicate that TBP contacts DNA as a monomer through interactions with the minor groove (18;19), causing a severe bend in the DNA around the TATA element (20) X-ray crystallographic structures of monomeric (60) and DNA-bound plant (63;64), yeast (65) and human (66) TBP have been described. All three reveal a very similar saddle-like structure which straddles the minor

groove of the DNA. The convex upper face of the saddle is lined with alpha-helices while the concave underside is a half-cylinder formed of beta-sheets and which is wide enough to accommodate a DNA molecule. Further, the contours of all three suggest the common induced-fit mechanism for protein-DNA recognition (reviewed in "22).

Upon cloning of the TATA box-binding component of TFIID, it was found that TBP could replace the TFIID fractions for basal levels of transcription. However, recombinant TBP failed to support *in vitro* activated transcription (see figure 2d) (49;54-67;69). This suggested that there is critical difference between TBP and TFIID. A "co-activator" hypothesis was suggested in which the TFIID chromatographic fraction must contain some further activities that are dispensable for basal transcription but essential for regulated transcription. Antibodies against TBP were used to affinity purify TFIID from human and *Drosophila* cell lines (67-74). It was found that nuclear TFIID from both human and *Drosophila* is composed of TBP and eight to twelve firmly associated TBP-associated factors (TAF_{II}s) (67;68;71-75). The apparent critical role in regulated transcription played by TAF_{II}s led to the gradual identification of their corresponding cDNAs. At present, there are some 15 distinct TAF_{II}s that have been cloned from human and *Drosophila*. These range in size 18 to 250 kDa and are highly conserved between the two species (see figure 2.4). Many of these TAF_{II}s also have interchangeable homologs (76) that have been identified in yeast, where they have been proven essential for viability (69;70;77;78). The diverse family of TAF_{II}s appear to have a complicated and equally diverse palette of functions, many of which overlap amongst combinations of TAF_{II}s. Such functions include regulatory interactions with activators and other GTFs, as well as enzymatic, cell-cycle, and nucleosome-like functions (see below; figure 2.4).

A simple nomenclature system has been adopted for the many TAF_{II}s that have been described in the three prominent organisms of transcription study. The suffix of each TAF_{II} is given as the factor's apparent molecular weight (in kDa) and the prefix "y", "h", or "d" corresponds to the genome from where it was isolated (yeast, human, or *Drosophila*).

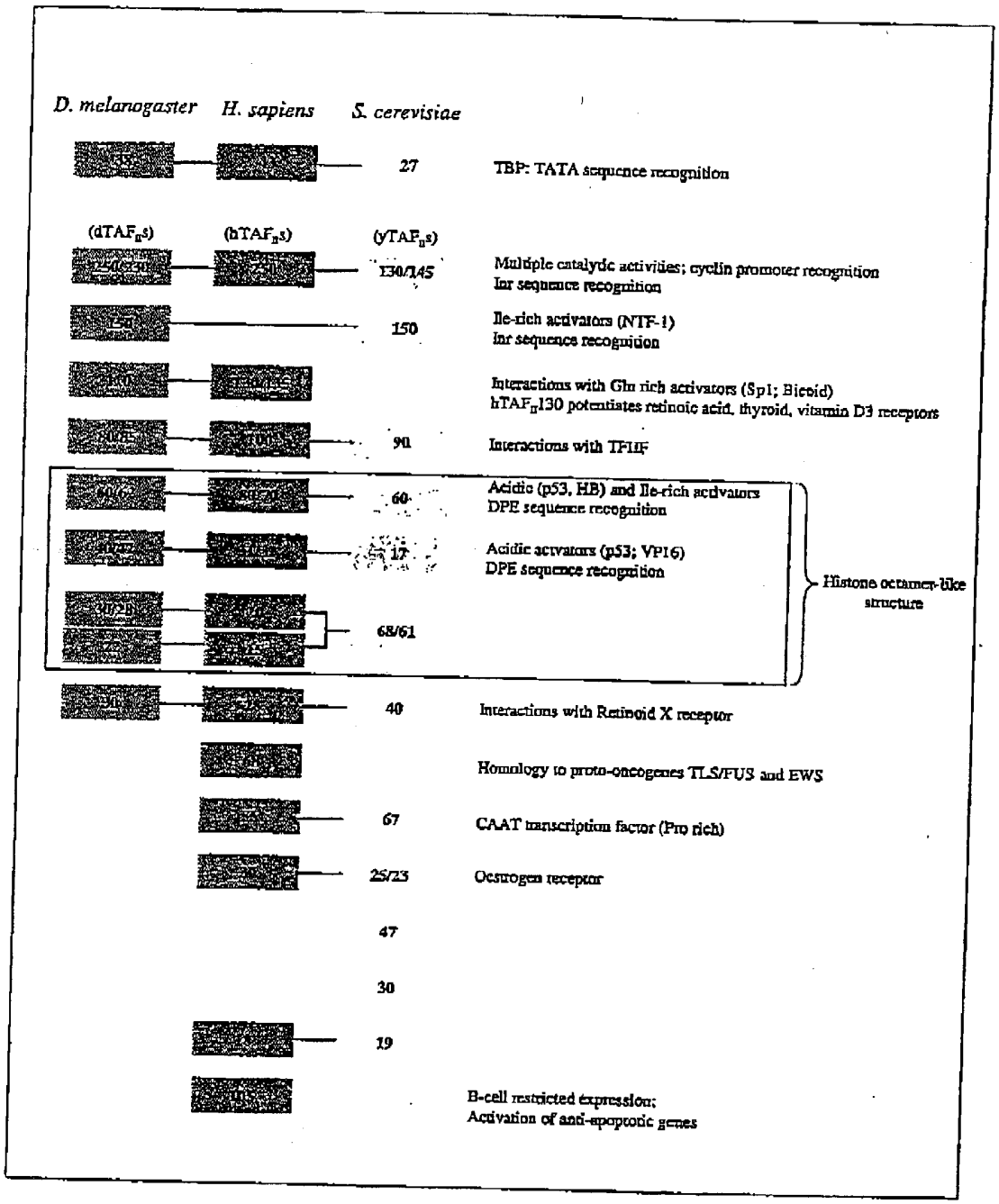


Figure 2.4: A comparison of TFIID subunits and their putative functions in *Drosophila*, human and yeast. Homologies are indicated with a line. Numbers indicate size in kDa. Some published size discrepancies are shown with "/ ". It should be noted that functions have not necessarily been observed experimentally in all members of a sequence-related group.

2.7 TAF_{II} functions: co-activators

The isolation of cDNAs for many individual TAF_{II} subunits permitted detailed biochemical analysis of the role of TFIID in transcriptional regulation. Consistent with the prior observation that the TFIID chromatographic fraction was a target for gene-specific activators (11;48), numerous activator-TAF_{II} interactions have been described in transcription studies performed *in vitro* using recombinant TFIID. It is now clear that different classes of activators require distinct subunits of the TFIID complex for activated transcription. For example, the $dTAF_{II}110$ and its human homolog $hTAF_{II}130$ enable the glutamine-rich activation domains of the enhancer-binding factors Sp1, Bicoid, and CREB by interacting selectively with them (79-81). By comparison, regulatory factors with acidic activation domains, such as p53 and VP-16, require selective interaction *in vitro* with $dTAF_{II}60$ and $dTAF_{II}40$ (82;83), or their human counterparts $hTAF_{II}70$ and $hTAF_{II}32$, to effect activated transcription. The largest TFIID subunit, $hTAF_{II}250$, appears to be critical for proper expression of cyclin genes by way of contacts with the ATF activator proteins, factors which interact with elements within cyclin promoters (84). Additional support for a "co-activator" function of TAF_{II}s has come from studies with various ligand-dependant nuclear receptors. A TFIID complex which contains human $TAF_{II}30$, a TAF_{II} factor which appears to have no *Drosophila* homolog, is required for activation of genes that are responsive to the oestrogen receptor (ER) (85). Exogenous expression of some TAF_{II} cDNAs in Hela and Cos cells has revealed that the human $TAF_{II}28$ is required for transcriptional activation through the retinoid X receptor (RXR)(86). Similar TAF_{II} transfection studies in human cell lines demonstrated that transcriptional activation through other ligand-dependant receptors, including the thyroid hormone, retinoic acid and vitamin D3 receptors, appears to be selectively potentiated through $hTAF_{II}130$ (87).

Much of the evidence for such activator-specific functions among the TAF_{II}s relies on reconstituted transcription reactions using recombinant components as well as artificial promoters transfected into immortal cell lines. A comprehensive *in vivo* observation of such a general co-activator role for TAF_{II}s has proven elusive. However, a recent study seems to strongly uphold the co-activator view *in vivo*, whereby the requirement of $dTAF_{II}110$ and $dTAF_{II}60$ for proper *dorsal*-mediated development in the *Drosophila* embryo is demonstrated (88).

2.8 TAF_{II} functions: TFIID recruitment and regulatory synergy

Although it is not clear exactly how such activation of transcription occurs, a simple interpretation is that activator-TAF_{II} contacts lead to an increased rate of recruitment of TFIID to the core promoter. An elegant *in vitro* study using partial TFIID, reconstituted from recombinant components, seems to support this view (89;90). The two enhancer-binding proteins Bicoid (BCD) and Hunchback (HB) are critical for correct expression of the pattern-forming genes in the developing *Drosophila* embryo. BCD and HB have been shown to activate transcription mutually exclusive of one another through interactions with $dTAF_{II}110$ and $dTAF_{II}60$, respectively (see above). With both factors present on the promoter, recombinant TFIID containing either $dTAF_{II}110$ or $dTAF_{II}60$ will exhibit modest transcriptional activation. When TFIID containing both $dTAF_{II}110$ and $dTAF_{II}60$ is used,

allowing simultaneous contact with HB and BCD, a greater-than-linear (synergistic) increase in transcriptional activation is observed. Promoter occupancy studies on identical *in vitro* transcription templates revealed parallel cooperative effects in recognition of the TATA box by TFIID (89;90). The minimum concentration of TFIID needed to achieve TATA box occupancy was much lower (~50 fold) when both BCD and HB were present than when only one or the other factor was present. This strongly suggests an activation mechanism that involves the recruitment of TFIID to the promoter through activator-TAF_{II} contacts.

2.9 TAF_{II} functions: core promoter selection

Besides their apparently critical co-activation role in regulated transcription, TAF_{II}s appear essential for a number of other transcriptional functions. As mentioned above, many promoters contain weak or even no TATA element. It has been shown that even basal levels of transcription are not possible from TATA-less core promoters in the absence of TAF_{II}s (91). Using recombinant TFIID, a combination of DNA binding studies and *in vitro* transcription assays have established the involvement of certain TAF_{II}s in discriminating between structurally distinct core promoters. The utilization of the Inr sequence elements in TATA-less promoters requires the minimal presence of $dTAF_{II}250$ and $dTAF_{II}150$ in the TFIID complex, with sequence specificity apparently being supplied soley by $dTAF_{II}150$ (92). Subsequent studies with $dTAF_{II}150$ revealed its partial role in mediating differential utilization of the two neighboring core promoters of the *Adh* gene in *Drosophila*. The developmental shift in transcription from the embryonic promoter to the adult promoter requires the presence of $dTAF_{II}150$ (93). Interestingly, several unidentified proteins are also reported to be involved which appear to be sub-stoichiometrically expressed $dTAF_{II}$s whose cDNAs have not yet been identified (93).

More recently, a novel core promoter element (consensus: a/gGa/tCTGT) has been identified which appears to be functionally analogous to the TATA-box and have a specific TAF_{II} interaction function. The "downstream promoter element" (DPE) appears ~30 base pairs downstream of the start site of many eukaryotic TATA-less core promoters (94). Further investigation has shown that $dTAF_{II}60$ and, to a lesser extent, $dTAF_{II}40$ can bring TFIID to TATA-less core promoters through direct interactions with a DPE, thus successfully supporting basal transcription with human or *Drosophila* GTFs (95).

It is generally understood that variation in the structure of promoters, such as the presence (or absence) of a potent or weak TATA-box, Inr sequence or DPE can vastly increase the range of potential mechanisms for transcriptional regulation. It appears that the distinct sequence-specific properties of TBP and certain TAF_{II}s, acting individually or in combination, allow TFIID to recognize a diverse array of core promoters, thus acting as a universal component of many regulatory mechanisms (reviewed in 12).

2.10 TAF_{II} functions: catalysis

In addition to its important interactions with regulatory factors and with promoter DNA, TFIID also possesses enzymatic activites. The TFIID subunit $TAF_{II}250$ ($TAF_{II}145$ in yeast) contains at least .two different catalytic activities; histone acetylation and GTF phosphorylation. Histone acetyl-transferase activity (HAT activity) is recognized to be an important step in the conversion of normally quiescent chromatin into a transcriptionally active form. It is thought that the acetylation of certain lysine residues in the histone complex weakens the histone-DNA interactions, thus enabling the transcription factor machinery access to the DNA. The ability of $TAF_{II}250$ to acetylate the H3, H4 and H2A histones has been conserved in yeast, *Drosophila* and human, likely indicating an important role for TFIID in controlling access to nucleosome-bound promoters *in vivo* (96). $TAF_{II}250$ is also contains two ATP-dependant serine/threonine kinase activities in separate domains, rendering it able to phosphorylate itself as well as RAP74, the large subunit of TFIIF (97). It is supposed that the rate at which the TFIIF-RNA Pol II complex is recruited to the PIC is sensitive to the phosphorylation state of RAP74. Studies have shown that disruption of the phosphorylase activity will impair transcription of at least two genes important for control of the cell cycle, cyclin A and cdc2, and can obstruct proper cell-cycle progression (98) . This suggests that this kinase activity is required for a specific subset of genes *in vivo*, and that such phosphorylation may be one of many modes of transcriptional regulation in which TFIID plays a pivotal role.

2.11 TAF_{II} functions: promoter topology

The TFIID subunits $dTAF_{II}30$, $dTAF_{II}40$ and $dTAF_{II}60$ and their human homologs are related in sequence and structure to the histone proteins H2B, H3 and H4, respectively (99;100). This raises evolutionary considerations for TAF_{II} origins as well as the intriguing possibility that parts of the TFIID complex may form a substructure similar to the histone octamer and which mediates TFIID-promoter interactions. Direct evidence for DNA wrapping around the TFIID complex in a histone-like fashion has been described (101). There is still little agreement about the exact role of such a histone-like structure within the TFIID complex. It has been proposed that it could represent a mode by which TFIID is involved in transcriptional regulation through the changes imposed by it on DNA topology, which may in turn induce conformational efficiency in GTF-promoter contacts (13;100). It has also been suggested that TFIID can mimic a bona fide nucleosome and actually become part of the chromatin packaging *in vivo*, perhaps thus becoming a "permanently activated" core promoter template for select, highly-expressed genes (102) .

2.12 Multiple TFIID complexes

In light of the wide repertoire of TAF_{II} functions that have been described to date, one of the most intriguing facets of TFIID biochemistry is the existence different TFIID subpopulations in the vertebrate nucleus. This possibility was first hinted at when it was noted that affinity purified TFIID

from human Hela cells appeared to comprise more subunits than that of *Drosophila* embryos (75;103). Laborious fractionation of Hela nuclear proteins revealed not only the existence of human TAF_{II}s with no apparent Drosophila counterpart (e.g. $hTAF_{II}30$), but also at least two functionally distinct TFIID complexes (called hTFIIDα and hTFIIDβ) in the same human nucleus (75). These discrete subpopulations of human TFIID have dissimilar abilities to support activated transcription (75;104). As hypothesized, closer examination revealed that the different hTFIID complexes contain different sets of TAF_{II}s. A "core" set of TAF_{II}s appears to be included in both hTFIIDα and hTFIIDβ, including $hTAF_{II}250$, $hTAF_{II}130$, $hTAF_{II}100$, and $hTAF_{II}28$ (85;105). Substoichiometrically expressed TAF_{II}s , whose cDNAs have since been isolated, accompany these core TAF_{II}s in hTFIIDβ, including $hTAF_{II}30$ $hTAF_{II}20$ and $hTAF_{II}18$ (85;106). Other low-abundance TAF_{II}s are reported to associate preferably with hTFIIDα. The corresponding cDNAs for these hTFIIDα-specific TAF_{II}s have not yet been isolated, but their apparent molecular weights (29, 64, 106 and 150 kDa) suggest they have no previously characterized human counterparts (85;106).

The implications of multiple TFIID complexes in the vertebrate nucleus are not trivial. As discussed above, it is clear that the numerous and disparate TAF_{II} subunits can bestow a diverse array of functions on this pivotal GTF. That the biochemistry of TFIID may include a combinatorial quality would vastly increase its functional versatility and, thus, its ability to act as a pivotal component in many regulatory mechanisms.

2.13 Tissue-specific TAF_{II}s

It is plausible that in metazoans, with highly elaborated tissue-specific and developmental programs of genetic expression, there would be a need for tissue-specific components of the core transcription machinery. The findings that certain human and *Drosophila* activators are unable to function in yeast (107-109) suggest the possibility that as eukaryotes evolved towards metazoans with more complex gene-regulatory pathways, there was a concomitant increase in the need for a more diverse network of co-activators. The reported combinatorial nature of TFIID offers the possibility that tissue-restricted TAF_{II}s may exist to fulfill this role (discussed above). However, the existence of such TAF_{II}s in metazoans would likely elude detection in the majority of current higher eukaryotic TFIID studies, which are based prevalently on Hela cell and *Drosophila* embryo nuclei. The use of Hela cells for such studies of human TFIID has been well established for over a decade. This is partly due to the fact that the human cervical carcinoma line need not grow on a two dimensional surface like many cells, but can be cultivated to extremely high density in liquid media culture. Therefore, like yeast and *Drosophila* embryos, Hela affords relatively economical access to large amounts of homogeneous source nuclei, a critical consideration when purifying low abundance nuclear proteins such as TFIID (110).

More recently, however, studies have begun to characterize TFIID in other human cell types. A novel 105 kDa $hTAF_{II}$ has been detected and characterized in a human B cell line (111), suggesting a

cell-type specific TFIID complex that could be responsible for mediating transcription by a subset of activators in B cells. Its transcript appears in other human cell lines, but at a much lower abundance (10 – 20 fold) (111). This suggests cell-type restriction of hTAF$_{II}$105 through substoichiometric distribution which is analogous to that of hTFIIDα and hTFIIDβ in the Hela nucleus. Further studies revealed that by serving as a co-activator for NF-κB through direct interaction, TAF$_{II}$105 is critical for the expression of anti-apoptotic genes (112). Although these studies originate from transformed cell lines, they further support the hypothesis of combinatorially distinct TFIID complexes residing in different nuclei of metazoan tissues.

Presently, hTAF$_{II}$105 is the only known TFIID subunit that can be considered to be tissue specific. However, it is widely held that in higher metazoans there likely exist more such TAF$_{II}$s which have escaped previous detection in the restricted milieu of Hela cell and *Drosophila* embryo nuclei. A promising observation comes from mRNA analysis of different rat tissue (brain, kidney, liver, lung, spleen, testes and uterus), revealing that the transcripts of many of the other known TAF$_{II}$s have multiple splice variants in different vertebrate tissues (113). This presents the possibility of multiple TAF$_{II}$s being generally transcribed from a unique gene locus, but following different tissue-specific mRNA processing pathways to produce tissue-restricted TFIID subunits.

2.14 Goals of this study

This study is intended to lay a novel groundwork for comprehensive investigation into tissue-specific aspects of transcription complexes, specifically, the TFIID complex. Based on previous TFIID studies in vertebrate cell lines, convincing arguments can be made to pursue investigation of holo-TFIID biochemistry in a whole organism. This would offer at least two advantages.

First, a whole animal system offers access to a full range of tissue types. Qualitative comparisons between tissues would be an invaluable tool for direct identification of putative new TAF$_{II}$s or similar factors. Moreover, fractions from different tissues could also be used to complement transcription of tissue-specific promoters *in vitro*, potentially giving insight into possible tissue-specific regulatory mechanisms.

The second advantage of using a whole organism is the ability to obtain transcriptional components from a wild-type background. Interpretations from studies using established cell lines suffer the shared caveat of an abnormality in the cell cycle machinery. This is unavoidable, since the inactivation of tumor-suppressor genes (cellular transformation) is recognized to be the *raison d'être* of any immortal cell line. A significant connection between the cell cycle and transcriptional regulation has long been recognized, since all transcription in the cell is arrested during mitosis and then resumed promptly afterward (reviewed in 114). More recent reports of crucial interactions between some components of the cell cycle and GTFs (albeit in Hela cells) are described elsewhere (reviewed in "115), including TBP and TFIID (116;117) *In vitro* studies using wild-type transcriptional holo-

components combined with select cell cycle-responsive promoters may reveal subtle features of the regulatory machinery that would otherwise be missed in studies using immortal cell line sources.

In this thesis work, the feasibility of using transgenically expressed, affinity-tagged TBP to co-purify associated functional elements of regulated transcription is demonstrated. In this whole-organism approach, we have created a novel mouse line that expresses a "double-tagged" version of TBP. As a central component of TFIID, tagged TBP accommodates efficient retrieval of the holo-complex from various tissues using a minimum of starting material. Two separate affinity tags are introduced onto the N-terminal primary sequence of TBP, with the intention of using multiple chromatography steps to yield a highly purified TFIID holocomplex. Similar strategies using a single N-terminal TBP tag have already been used in transfected cell lines to successfully isolate functional TFIID, and thereby, many individual TAF_{II}s (discussed above). Although some groups have developed elaborate purification schemes to isolate GTFs from select organ material, most notably rat liver (118), a transgenic purification approach has not been used previously in any metazoan system.

3. MATERIALS AND METHODS

3.1 Materials

3.1.1 General equipment

DNA Thermal Cycler	Perkin Elmer
Phosphor-Imager	BioRad
Horizontal agarose slab-gel apparatus	Hoefer-Pharmacia HE33 and HE99
Polyacrylamide Gel Apparatus (10x12cm)	Hoefer-Pharmiacia SE295
Polyacrylamide Gel Apparatus (18x16cm)	Hoefer-Pharmiacia SE600
Homogenizer-(Mikrodismembrator)	Braun-Melsungen
Peristaltic Chromatography Pump	Gilson
Semi-dry electroblotter	Hoefer-Pharmacia
Speed-Vac	Eppendorf
"Milli-Q Ultra Pure" water filter	Millipore
Oligonucleotide synthesis machine	MWG Biotech
Spectrophotometer	LKB
Ultra-Turrax T25 tissue homogenizer	Janke & Kunkel
Tabletop microcentrifuge 5415C	Eppendorf
NucTrap probe purification columns	Stratagene

3.1.2 General materials

PVDF Membrane	Millipore
Talon metal affinity resin	Clontech
P11 Phosphocellulose resin	Whatman
Protein assay reagent	Bio-Rad
Non-fat dehydrated milk	Bio-Rad
Chemilluminescent Western Blotting Substrate	Roche-Boehringer-Mannheim
Glass Chromatography Columns	Bio-Rad (Econo-columns)
Size exclusion columns ("Push Columns")	Stratagene
3MM chromatography/blotting paper	Whatman

3.1.3 General and specialized chemicals

All chemicals are "analysis" or "molecular biology" grade chemicals from Sigma, Merck CalBiochem, or Riedel-de Haen.

3.1.4 Commercially available kits

Plasmid "maxi-prep" kit	Qiagen
Plasmid "mini-prep" kit	Qiagen
Site-directed mutagenesis kit	Clontech
TnT coupled reticulocyte system	Promega
Silver staining kit	Pharmacia

3.1.5 Enzymes

Alkaline phosphatase	Roche-Boehringer-Mannheim
Restriction enzymes	Roche-Boehringer-Mannheim

Taq DNA Polymerase	Perkin-Elmer
T4 DNA Ligase	Roche-Boehringer Mannheim
S1 Nuclease	Gibco/BRL

3.1.6 Radioisotopes

$\gamma[^{32}P]$ ATP	3000 cpm/mmol	Amersham
$\alpha[^{32}P]$ CTP	3000 cpm/mmol	Amersham

3.1.7 General buffers and solutions

Milli-Q water: Sterile and deionized water that was treated by passage through a Milli-Q Ultra-Pure water filter (Millipore).

10x TE buffer: 10mM Tris-HCL (pH 8.0)
1mM EDTA

50x TAE: 2M Tris-HCl
57.1ml glacial acetic acid (per liter)
50mM EDTA
(final pH ~ 8.5)

10x TBE buffer: 1M Tris
1M Boric Acid
20mM EDTA

BC buffer: 20mM Tris-HCl (pH 7.3 at 20°C)
0-1000mM KCl
20% Glycerol
5mM DTT (added just prior to use)

3.1.8 Bacterial strains

Bacterial strain	*Purpose*	*Genotype*
HB101	plasmid growth	F- Δ(gpt-proA)62 leuB6 supE44 ara-14 galK2 lacY1 Δ(mcrC-mrr) rpsL20(str^r) xly-5 mtl-1 recA13
BMH 71-18 mutS	mutagenesis	thi supE Δ(lac-proAB) [mutS::Tn10][F' proAB, lacI^q ZΔM15]

3.1.9 Plasmids

Plasmid	*Function*	*Source*
PAG-17	Vector containing double tagged hTBP cDNA	M. Meisterernst
PPBV	Cloning of metallothionine promoter-driven transgene	Pharmacia
PEFBG	Cloning of EF-1α promoter-driven transgene	G. Polites
PMRG5	Template for in vitro transcription	M. Meisterernst
PMLΔ53	Template for in vitro transcription	M. Meisterernst

3.1.10 DNA length standards

Standard	*DNA fragment lengths in base pairs (bp)*
Kilobase DNA ladder (Gibco/BRL)	12216, 11189, 10180, 9162, 8144, 7126, 6108, 5090, 4072, 3054, 2036, 1636, 1018, 517, 506, 396, 344, 298, 220, 201, 154, 134, 75
DNA Marker V (Roche-Boehringer-Mannheim)	587, 540, 504, 458, 434, 267, 234, 213, 192, 184, 124, 123, 104, 89, 80, 64, 57, 51, 21, 18, 11, 8
HindIII digested λ-phage DNA (Gibco/BRL)	23130, 9416, 6557, 4361, 2322, 2027, 564, 125

3.1.11 Oligodeoxyribnucleotides

Oligo	*Purpose*	*Sequence (5'→3')*
MT(-150)	PCR	GGAGC AACCG CCTGC TGGGT GC
EF(-70)	PCR	GGAGA CTGAA GTTAG GCCAG C
TBP(+430)	PCR	CCTGT GTTGC CTGCT GGGAC G
RI/RV	mutagenesis	GCATC CCGGA TATCC TGCAG CCCAA CATGG CCGCT TCG
S1/5'	S1 nuclease protection	GCGGC ACCAG GCCGC TGCTG TGATG ATGAT GATGA TGGCT GCTGC CCATG ACTGC GTAAT GCGGT CATGA CGCTT T
S1/3'	S1 nuclease protection	GAAGG GGGTG GGGGA GGCAA GGGTA CTGAG AGCCA TTACG TCGTC TTCCT GAATC CCTTT AGCCG CTTTG CTCG
Xba-Not/fwd	cloning adapter	CTAGA GAACG TCAC
Xba-Not/rev	cloning adapter	GCGGC CGTGA CGTTC T
Xho-Nco/fwd	cloning adapter	TCGAG AACGT CTCC
Xho-Nco/rev	cloning adapter	CATGG GAGAC GTTC

3.1.12 Protein size standards for SDS-PAGE

SDS-PAGE Rainbow Protein Standard low range — Amersham
SDS-PAGE Rainbow Protein Standard high range — Amersham

Protein	*MW (Da)*	*Low range*	*High range*
Myosin	200,000		x
Phosphorylase B	97,400		x
Bovine Serum Albumin	66,000		x
Ovalbumin	46,000	x	x
Carbonic anhydrase	30,000	x	x
Trypsin Inhibitor	21,500	x	x
Lysozyme	14,300	x	x
Aprotinin	6,500	x	
Insulin ($\alpha + \beta$ chain)	3,400 + 2,350	x	

3.1.13 Antibodies

Anti TBP 58C9 mouse mAb — Santa Cruz Biotech
Anti HA rat mAb "high affinity" — Roche-Boehringer Mannheim
Sheep Anti-rat (Peroxidase-linked) — Amersham
Sheep Anti-mouse (Peroxidase-linked) — Amersham

3.2 Methods

3.2.1 DNA Manipulation

3.2.1.1 Determination of nucleic acid concentration and purity

To determine the concentration of DNA, RNA, and oligonucleotide preparations, the spectrophotometric absorption of the aqueous solutions at 260nm and 280nm was measured. For DNA, 1 OD unit at 260nm corresponds to a concentration of 50µg/ml. For RNA, 1 OD unit at 260nm corresponds to 40µg/ml. For oligonucleotides (single stranded DNA less than 100 bases in length), 1 OD unit at 260nm corresponds to 20µg/ml.

The ratio between the readings at 260 and 280nm (A_{260}/A_{280}) provided an estimation of the purity of the nucleic acid. Pure preparations of DNA and RNA will have an A_{260}/A_{280} ratio between 1.8 and 2.0. Ratios below 1.8 indicate impurities in the preparation (protein, phenol, lipids, etc).

3.2.1.2 Precipitation of nucleic acids

In general, nucleic acids (>1µg) in small aqueous volumes can be precipitated in ethanol using sodium acetate. 1/10 volume of 3M sodium acetate was added to the DNA-containing solution, followed by 2.5 volume of 96% ethanol. The solution was then allowed to precipitate at least 1 hour at –20°C (or 15 minutes at -80°C). DNA was then pelleted by centrifuging at 14,000rpm (16,000xg) for 5 minutes. The aqueous solution is then removed, leaving a DNA-salt pellet. Excess salt is then washed from the pellet by adding 75% ethanol (pre-chilled to –20°C) and re-centrifuging. Wash steps to remove salt were usually carried out no more than twice. Pelleted DNA can then be air dried and re-dissolved in an appropriate buffer and volume.

3.2.1.3 Purification of oligonucleotides

Oligonucleotides were synthesized on the premises by B. Eidenmueller using an oligonucleotide synthesizer from MWG. Remnants of chemical synthesis ("protecting groups") were removed from the finished oligonucleotides with heat treatment (65°C for 90 minutes) followed by desiccation in a Speed-Vac. Oligonucleotides were then redissolved in 1xTE/10mM $MgSO_4$ and precipitated at -20°C with 5 volumes of 96% ethanol. After washing at least once with 70% ethanol, the oligonucleotides were resuspended in TE and quantitated (described in section 3.2.1.1). Longer oligonucleotides (> 20 bases) were further purified with denaturing polyacrylamide gel electrophoresis as described in chapter V.2.12 of *Protocols in Molecular Biology* (119).

3.2.1.4 Hybridization of oligonucleotides to create cloning linkers

Equal amounts of each single stranded oligonucleotide (25-50μg) were combined in 50-100μl total volume of 1x PCR reaction buffer (Perkin-Elmer). Oligonucleotide mixture was then treated at 60°C for 15 minutes in a water bath to ensure complete dissociation. The water bath was then turned off and allowed to gradually cool to room temperature (3 hours to overnight). Hybridized linkers were then precipitated with 0.3M sodium acetate (described in section 3.2.1.2) and washed.

3.2.1.5 Restriction enzyme digestion of DNA

DNA used for construction of transgenes was digested with restriction enzymes and accompanying buffers from Roche-Boehringer-Mannheim. Generally, the reactions consisted of 1 to 50μg of DNA in final reaction volumes of 20 to 100μl. Individual reaction conditions for each enzyme (activity units of enzyme, temperature, additional buffer components, etc), as stated by the manufacturer, were followed.

3.2.1.6 Insertion of DNA fragments into plasmids

Construction of plasmids containing the entire transgenic construct was carried out by ligating the cDNA fragment into plasmids vectors containing the correct promoter and 3' polyadenylation signal. Fragment and vector were produced by cutting the parent plasmids with appropriate restriction enzymes. Insert DNA was resolved by size from its parent plasmid on a 0.7% agarose gel. The fragment could then be cut out of the gel, and the DNA purified from the agarose matrix using phenol extraction. In cases where the restriction enzyme-generated overhangs of the vector and insert did not match, short adapter DNA was synthesized. An adapter is created from two synthesized oligonucleotides sequences which have been hybridized to one another, yielding a double stranded DNA with specific single-stranded end overhangs. This facilitates the ligation of two otherwise non-matching DNA overhang ends.

For ligation reactions, 50-200ng of cut vector plasmid was combined with insert and adapter in 30μl reactions. The reaction volumes also contained 1x T4 DNA ligase buffer and 200-400 units of T4 DNA ligase (both from Roche-Boehringer-Mannheim), and 1mM ATP. The amount of insert and adapter added was in molar proportionality to the amount of vector. Generally, 3 different reaction ratios (vector:insert:adapter) were attempted for each ligation; 1:2:10, 1:5:10, and 1:10:10. Reactions were allowed to proceed for 16 hours at 16°C.

3.2.1.7 Mutagenesis of plasmid

To make one of the DNA constructs for production of transgenic mice, it was necessary to abolish an EcoRI site at the 3' end of the gene for tagged hTBP (see figure 4.2). This was accomplished with a site-directed mutagenesis kit (Clontech) according to the instructions therein. This kit is based on the method of Deng and Nickoloff (120), in which specific base changes may be introduced into any

double stranded plasmid. The procedure relies on the existence of at least one unique restriction enzyme site on the plasmid which is abolished in the course of introducing a specific desired mutation elsewhere on the plasmid. Normally, two oligonucleotide primers are needed; one for the desired mutation and one for selection through elimination the unique restriction site. In the case of the mutation introduced into pAG-17, the site of the desired mutation and the unique restriction site were very close to each other. Therefore it was possible to design a single oligonucleotide for introduction of both changes simultaneously.

3.2.1.8 Generation of competent *E. coli*

A flask containing 50-100ml of LB was inoculated with 0.5-1.0ml of a stationary phase culture of HB101. This was allowed to grow at 37°C on a shaker for 2-3 hours until the OD at 600nm was between 0.2 and 0.4. The culture was then placed on ice for 10 minutes, then transferred to 50ml sterile centrifuge tubes and pelleted at 3000rpm for 10 minutes. The pelleted bacteria were then gently resuspended in sterile, ice cold 0.1M calcium chloride (10 ml for each 50 ml of original culture volume). This mixture was allowed to incubate on ice for 30 minutes then centrifuged again at 2500 rpm at 4°C for 10 minutes. The bacterial pellet was then carefully resuspended in sterile, ice-cold 0.1M calcium chloride (1ml for each 50ml of original culture volume). This condensed volume of bacteria was allowed to incubate 30 minutes longer on ice. Competent bacteria could then be used for transformation of plasmids from purified preparations or from ligation reactions.

3.2.1.9 Transformation of DNA into competent *E. coli*

In a sterile 1.5ml Eppendorf reaction tube, 50-100µl of competent cells were combined with 1-200ng of purified plasmid or 25% of a ligation reaction (5-10µl). In all transformations, the plasmids contained the gene for ampicillin resistance (β-lactamase) The mixture was incubated on ice for at least 30 minutes. The mixture was then heat shocked at 37°C for exactly three minutes, then incubated further on ice for 30 minutes. Under aseptic conditions, the mixture was then transferred to 5ml sterile LB and allowed to incubate shaking at 37°C for 30-60 minutes. The bacteria were then pelleted at 3000rpm for 10 minutes. The pellet was then transferred evenly onto LB agar-ampicillin plates and allowed to incubate for 16 hours at 37°C. The resulting bacterial colonies originate from bacteria which have been transformed with plasmid containing β-lactamase.

3.2.1.10 Growth of plasmid-containing bacterial cultures

Plasmids utilized in this work contained the β-lactamase gene. Cultures of transformed bacteria (HB101) were therefore ampicillin resistant (*amp*^{r}) and could be selectively grown in ampicillin-containing medium. Generally, a volume of LB-Amp media was inoculated in a sterile environment under with a small amount of *amp*^{r} bacteria. Cultures were allowed to grow overnight at 37°C while

shaking for aeration. In some cases (e.g. ligations), it is necessary to grow bacteria as colonies on a solid surface. In this case, plastic petri-style dishes containing LB-agar plates were used.

LB medium:	10g/liter casein hydrolysate (Bacto-tryptone) 5g/liter yeast extract 10g/liter NaCl pH adjusted to 7.0 with NaOH, then autoclaved.
LB plates:	LB medium prepared with 1.5% agar. After melting the agar by autoclaving and cooling to ~45°C, the LB agar is poured into agar plates under sterile conditions.
LB-amp:	Solid ampicillin was added as required to LB medium and LB-agar after autoclaving and cooling to ~45°C. Final concentration: 50µg/ml.

3.2.1.11 Small scale plasmid DNA "mini" preparation

Rapid isolation of plasmid DNA from single or multiple bacterial cultures was carried out with the "QIAprep 8 Plasmid Kit" (Qiagen). This kit is based on the same plasmid purification principle (anion exchange matrix) as the Qiagen Plasmid Kit, using rather an abbreviated protocol and smaller columns. The kit includes specialized components which accommodate simultaneous preparation of up to 48 different plasmids. The exact protocol was carried out exactly as described in the accompanying "QIAprep 8 Plasmid Kit" handbook for 5ml bacterial cultures. All lysis and washing buffers were also included with the kit. Purified plasmid DNA was stored in 1xTE at -20°C.

3.2.1.12 Large scale plasmid DNA "maxi" preparation

Generation of preparative amounts of purified plasmid DNA (0.2 – 1mg) was facilitated with the "Qiagen Plasmid Kit" (Qiagen). This method is based on lysis of the bacterial membrane using alkali conditions (KOH), thus releasing total bacterial DNA (genomic and plasmid). The genomic DNA stays primarily associated the discarded membrane. Released plasmid DNA is bound to an anion exchange column which is supplied with the kit. The bound plasmid can be washed of bacterial contaminants (proteins, lipids, etc), then removed by changing the salt concentration and pH of the column buffer. The exact protocol was carried out exactly as described in the accompanying "Qiagen Plasmid Kit" handbook for 500ml bacterial cultures. All lysis and washing buffers were also included with the kit. Purified plasmid DNA was stored in 1xTE at -20°C.

3.2.1.13 Native agarose gel electrophoresis

Agarose gel electrophoresis was carried out in a horizontal slab-style apparatus (Pharmacia-Hoefer) in two different sizes. The Hoefer HE 33 (8cm x 5cm) is the smaller of the two, normally accommodating up to eight sample lanes. The Hoefer HE 99 (20cm x 16cm) can accommodate one or two rows of up to 20 lanes. Depending on the size(s) of DNA to be resolved, gels containing 0.7-1.5% (w/v) agarose were used. To make a gel, the corresponding amount of solid agarose was added to 1x

TAE (30 ml total volume for HE 33; 200ml for HE 99). The solid agarose dissolved upon heating, whereupon a magnetic stirrer was added to facilitate even cooling of the agarose solution. When the temperature of the solution had decreased to at least 50°C, 1µl of a 10mg/ml solution of ethidium bromide was added in order to later visualize the DNA with UV light. When fully mixed, the melted agarose solution could be poured into the sealed HE 33 or HE 99 gel tray. The comb form was then added and the gel was allowed to solidify at room temperature. The solid agarose matrix and the platinum electrodes could then be covered with an uninterrupted volume of 1x TAE. 10-15% (v/v) DNA gel loading buffer was then mixed into to the DNA sample before being pipetted into the bottom of each sample well. Depending on the length of DNA to be resolved and the density of the matrix, the gel was run at 25-110V until the bromphenol blue had migrated 25-100% through the length of the gel bed. Ethidium bromide stained bands could be visualized and photographed by laying the gel on a UV light box.

50x TAE:	2M Tris-HCl 57.1ml glacial acetic acid (per liter) 50mM EDTA (final pH ~ 8.5)
10x DNA gel loading buffer:	10x TAE 30% glycerol 1mM EDTA 0.25% bromphenol blue 0.25% xylene cyanol

3.2.1.14 Elution of DNA out of agarose gel

Plasmid DNA that had been digested with appropriate restriction enzymes was size resolved on a 0.7% agarose gel (described in section 3.2.1.13). Using a UV light to visualize the ethidium bromide stained DNA bands, the desired band was identified based on size and removed from the gel by cutting a small piece out with a razor blade. The piece of solid agarose was chopped as finely as possible with the razor and transferred into a 1.5ml Eppendorf tube. 750µl Tris-phenol was added and the solution was vortexed. The tube was then allowed to freeze for at least 30 minutes at -80°C, followed by centrifugation at 14,000rpm (16,000xg) for 10 minutes. The aqueous upper phase was removed and 500-750µl 1xTE was added. The same vortexing, freezing and centrifuging steps were repeated and the aqueous phase was removed and combined with the first. An equivalent volume of Tris-phenol-chloroform was then added and the solution was vortexed and centrifuged again. The aqueous phase was removed again and the fragment DNA precipitated (described in section 3.2.1.2).

Tris-phenol:	Phenol is repeatedly mixed with equal volumes of 1M Tris-HCl (pH 7.2), allowed to phase separate, and then removed and mixed with fresh Tris-HCl until the pH of the phenol phase increases to ~7.
Tris-phenol-chloroform:	Phenol is mixed 1:1 with chloroform and the mixture is saturated with Tris as described above.

3.2.2 Creation and testing of transgenic mice

3.2.2.1 Microinjection and breeding

The microinjection of transgenic DNA into fertilized mouse oocytes and subsequent generation of transgenic mouse lines was carried out in the laboratory of H. Greg Polites at Hoechst-Marion-Roussel (Bridgewater). The initial injections were carried out using F1 hybrid mice from C57BL/6 x SJL litters. Later, the lines were bred into Balb/c mice as well. All of the mice in this work were heterozygous for the transgene. Details of the standard microinjection techniques that were used can be found in Chapter 2 of *Transgenic Animal Technology: A Laboratory Handbook* (121).

3.2.2.2 Extraction of genomic DNA from newborn mouse tails

Genomic DNA of possible transgenic mice was extracted from a small piece of tail tissue (~ 1 cm) cut from 2-4 week old offspring. Tail sections were incubated overnight on a shaker in 500-750µl tail buffer at 54°C. An equal volume of phenol/chloroform/isoamyl alcohol (25:24:1) was then added to each sample, which was then shaken gently for 15 minutes by hand or by vortexing at a low setting. Tail tissue samples were then centrifuged for 10 minutes at 14,000rpm (16,000xg). The upper aqueous phase was transferred to a 5ml polypropylene tube and an equivalent volume of 95% ethanol was then added drop wise to the tube, allowing the DNA to gradually precipitate. A second equivalent volume of 95% ethanol was then added more forcefully in order to mix the contents of the tube. Tubes were then inverted several times to ensure thorough mixing. Visibly precipitated genomic DNA from each tail sample was spooled around a flat pipette-tip and transferred to an individual well of a 96-well microtiter dish. DNA was air dried at ambient temperature in a for 4 hours. 200µl of 1x TE was then added to each well, and genomic DNA was then allowed to dissolve for 15-30 minutes at room temperature. Gentle pipetting up and down ensured complete mixing. Microtiter plates containing genomic DNA from multiple mice were normally sealed with paraffin and stored at -20°C until needed.

Tail buffer: 10mM Tris pH 7.5
100mM NaCl
10mM EDTA
0.5% SDS
30µg/ml Protease K (Sigma)

3.2.2.3 PCR testing of genomic tail DNA

DNA isolated from tail tissue was first thawed to room temperature. 10µl of each sample transferred to an identical 96-well plate and digested with EcoRI enzyme for at least 2 hours. The digested genomic DNA was then diluted 1:3 with water, and the entire plate was covered and heated to 95°C for 10 minutes. 2µl of each sample was used for PCR reactions. PCR was performed using AmpliTaq enzyme and buffer (Perkin-Elmer). Reactions were carried out in a 50µl volume consisting of

2µl genomic DNA, 1x reaction buffer, 1.5mM $MgCl_2$, 0.8µM forward oligonucleotide primer, 0.8µM reverse oligonucleotide primer, 8µl of nucleotide mixture containing 0.2mM of each nucleotide (dATP, dCTP, dTTP, and dGTP), and 2.5 units of AmpliTaq enzyme. Temperature cycling for *taq* amplification was carried out using a the following temperature cycle profile

Step 1; DNA melting for 2 minutes at 95°C
Step 2: Primer hybridization for 1minute at 55°C
Step 3: *Taq*-mediated DNA synthesis for 1 minute at 72°C

Cycling of steps 1 through 3 was repeated 40 times, followed by storage at 4°C. The presence of amplified product was detected by running standard agarose gel electrophoresis (1.2% agarose) and staining the gel with ethidium bromide for visualization of DNA with UV light (described in section 3.2.1.13).

3.2.2.4 Induction of transgenic promoter in MT-hTBP mice

Prior to RNA or protein isolation it was necessary to induce transgenic expression in the MT-TBP mice, since the MT-1 promoter is active only in the presence of Zn^{2+}. For each MT-TBP mouse, an interperitoneal injection of $ZnSO_4$ in sterile water was made 4 hours prior to animal sacrifice using a thin gauge syringe needle (#27). 0.1mg $ZnSO_4$ was injected per 10g mouse weight. The injected volume was always between 200µl and 400µl.

3.2.2.5 DEPC treatment of buffers and water

Water and buffers for RNA preparation were treated with DEPC prior to use. DEPC inactivates RNAse contaminants by covalent modification. Pure DEPC was added to milli-Q water and buffers to a final concentration of 0.002% (v/v). The flasks were then shaken vigorously and allowed to sit for at least 2 hours. Solutions were autoclaved afterward, ensuring complete breakdown of DEPC.

3.2.2.6 Isolation of total RNA from mouse tissue

Total RNA was extracted from 0.5-1g of mouse liver, brain, kidney, lung or muscle using commercially available Trizol reagent (Gibco-BRL). In a 14ml polypropylene tube (Falcon #2059), fresh or thawed tissue was homogenized with an Ultra-Turex tissue-homogenizer in 5ml of Trizol solution. After adding 1ml of chloroform, the tubes were capped and shaken vigorously by hand for 15 seconds. The tubes were then incubated at ambient temperature for 3 minutes, followed by centrifugation at 12,000xg for 15 minutes at 4°C in an SS-34 rotor (Sorvall-DuPont). The aqueous supernatant was removed to a new polypropylene tube and 2.5ml 100% isopropanol was added and mixed by inversion. Samples were then left at ambient temperanre for at least 10 minutes, allowing RNA to precipitate. RNA was then pelleted by centrifugation at 12,000xg for 10 minutes at 4°C. The supernatant was then decanted and the RNA pellet washed with 5ml 75% ethanol and re-centrifuged at

7,500xg for 10 minutes. The final RNA pellet was resuspended in DEPC-treated water and quantified (described in section 3.2.1.1). RNA preparations were routinely stored for later use at -80°C.

3.2.2.7 5' end-labeling of oligonucleotides and DNA marker with ^{32}P

40ng of oligonucleotide or 1µg of DNA marker V was 5'-labeled with [γ-^{32}P]-ATP (3000cpm/mmol) using T4 polynucleotide kinase and 1x buffer from Roche-Boehringer-Mannheim in a final reaction volume of 50µl. Labeling reaction was allowed to run for 1 hour at 30°C. Labeled oligonucleotide was separated from unincorporated ^{32}P using NucTrap columns (Stratagene). The columns were used exactly as recommended by the manufacturer.

3.2.2.8 Detection of transgenic mRNA by S1 nuclease protection

All solutions were pretreated with DEPC to ensure deactivation of RNAse. Uniform amounts of total RNA (10-20µg) were brought up to 100µl with DEPC-treated water in 1.5ml eppendorf tubes. 100,000 cpm of labeled oligonucleotide (0.1-1ng, depending on labeling efficiency) was added to each RNA sample. The entire RNA-oligonucleotide mixture was then precipitated with 0.1 volume 3M sodium acetate followed by 5 volumes of 95% ethanol. RNA and oligonucleotide were pelleted at 14,000rpm (16,000xg) for 5 minutes at 4°C. The pellet was carefully washed once with 75% ethanol, then air dried at ambient temperature. RNA and labeled oligonucleotide were then resuspended in 23µl S1 hybridization solution. The mixture was then denatured at 65°C for 20 minutes. 5M NaCl was preheated to 65°C and then 2µl was added quickly to each sample tube. RNA samples were incubated 1 hour further at 65°C in a water bath, then the temperature of the water bath was reset to 37°C. The gradual decrease in temperature from 65°C to 37°C (overnight) facilitated specific hybridization of anti-sense oligonucleotide to matching mRNA. The next day, 300µl of freshly prepared S1 reaction buffer was added to each sample. Samples incubated at room temperature for one hour, then 1ml 95% ethanol (pre-chilled to -20°C) was added to precipitate all nucleic acid. Centrifugation at 14,000xg for 10 minutes at 4°C pelleted all precipitated nucleic acid. The pellet was washed 3 times with 75% chilled ethanol, then dried briefly in Speed-Vac. It was important that the RNA pellet not become completely desiccated in the Speed-Vac, therefore only about 5 minutes was needed. The semi-dried pellet was then resuspended in 12µl S1 loading buffer. Samples were then heated to 70°C for 5 minutes before being loaded and size separated using denaturing (7M urea) thin polyacrylamide gel electrophoresis.

S1 hybridization solution:
80% formamide
10mM PIPES (pH 6.4)
1mM EDTA
0.05% SDS

S1 reaction buffer:	167 units/ml S1 nuclease (Gibco) 0.3M NaCl 30mM sodium acetate (pH 4.5) 3mM $ZnSO_4$
S1 loading buffer	1x TBE 85% formamide 0.01% bromphenol blue 0.01% xylene cyanol

3.2.2.9 Denaturing gel electrophoresis for S1 nuclease analysis

For resolution of protected oligonucleotides after S1 nuclease digestion, "sequencing style" denaturing gel electrophoresis was used. The sequencing gel system from Gibco-BRL was utilized with 60x40cm glass plates. Before constructing the gel, the front and back glass plates were washed extensively with SDS, then briefly with acetone and isopropanol. The inner surface of the front plate was subsequently siliconized with SigmaCote solution (Sigma). The plate sandwich was assembled with lightly greased 0.4mm spacers and clamped together. The 6% acrylamide (1:19 cross linking) gel mix was prepared with 7M urea and quickly injected in between the plates after adding 1% (v/v) ammonium persulfate (a freshly made 10% solution) and 0.1% (v/v) TEMED. The top of the gel was sealed off with the blunt side of a "shark tooth" comb during polymerization. It was then removed and reinserted with the tooth points slightly entering (~1mm) the acrylamide matrix, thus forming the sample wells.

The glass plate and gel assembly was then fitted into the vertical stand, and the upper and lower electrode chambers filled with 1xTBE. The gel was pre-run at 1800V for 30 minutes to warm the gel to ~50°C, thus hindering the formation of secondary structure in loaded samples. Prior to loading the gel, excess urea was rinsed from each sample well with a thin jet of 1xTBE. The S1 nuclease analysis samples were run at 1800V until the bromphenol blue dye had migrated through 2/3 of the gel length. The gel sandwich was then cooled, carefully opened, and the 0.4mm gel gently transferred to 3MM Whatman chromatography paper. After covering with thin plastic wrap, the gel was dried under vacuum at 80°C. Bands of ^{32}P-labeled oligonucleotides were detected by exposure to film (autoradiography) or by use of a Phosphor-Imager device (BioRad).

3.2.2.10 Preparation of nuclear extract from liver, brain and kidney

Mice were sacrificed by cervical dislocation and brain, kidney and brain tissue were removed. Mice carrying the MT promoter were induced four hours prior to sacrifice as described in section 3.2.2.4. Tissue could be used immediately for production of nuclear extracts, but was normally flash frozen in liquid nitrogen, and then stored at -80°C.

For nuclear extract production, frozen tissue was put in a solid steel mortar and pestle, and then crushed into small pieces with an ordinary hammer. The mortar and pestle were partially submersed in a pool of liquid nitrogen to maintain low temperature conditions. Crushed tissue was then transferred into the 25ml teflon homogenization capsule of the Micro-Dismembrator (Braun-Melsungen). The entire

teflon assembly was also pre-cooled by submersion in liquid nitrogen prior to transfer of crushed tissue. The capsule was filled about 50%, and two pre-chilled 0.8cm steel pellets were added. The assembly was closed and submersed in liquid nitrogen for at least 30 seconds before attaching to the Micro-Dismembrator and shaking at the highest setting for 20-30 seconds. The capsule was then quickly removed and re-submerged in liquid nitrogen. This homogenizing and re-cooling cycle was repeated 3 - 4 times. Maintaining the tissue at a very low temperature gave a thoroughly homogenized tissue which had a consistency of fine powder. Since the teflon capsule has limited capacity, homogenized tissue was accumulated in a 50ml polypropylene tube (Falcon #2073) on dry ice.

When all tissue for a nuclear extract preparation had been successively crushed, then homogenized, the accumulated tissue was allowed to warm gradually on normal ice. As the powdered tissue thawed, 15-20ml of ice cold homogenization buffer was added to each tube. The thawing tissue and buffer solution was then further homogenized with an Ultra-Turrax T25 tissue homogenizer. The device was set at 1/3 full power and each tube was homogenized for 3-4 times for 30 seconds. Since the friction of the homogenizer can warm the solution, efforts were taken to ensure that the tubes were always on ice in between steps. Afterwards, each of the tubes was brought up to the full volume (50 ml) with ice cold homogenization buffer. The total amount of tissue used was 6-10g/50 ml homogenate. Into the bottom of an appropriate number of 25x89mm polystyrene "UltraClear" centrifuge tubes (Beckman), 8 ml of fresh, ice cold homogenization buffer was added. To each of these tubes was added 25ml of tissue homogenate. The viscous tissue homogenate was added gently so as not to mix with the fresh buffer, but rather to form a second phase on top of it. The tubes were then carefully transferred to a pre-chilled SW28 swing-bucket centrifuge rotor, and centrifuged at 25,000 rpm at 4°C for 1 hour.

After centrifugation, a compact pellet of nuclei was apparent at the very bottom of the tube in the clear phase. The upper aqueous and lipid phases were carefully removed so as not to disturb or contaminate the pellet. Depending on the size of the pellet, 400 to 800µl of ice cold NexB 0.02 was added. The pellet was gently dispersed into solution with a small magnetic stir bar while submersed in an ice bath. When fully dispersed, an equi-volume of chilled 0.6 NexB was added drop wise over 30-40 minute period while continuously mixing with the stir bar (final [KCL] ≈ 300mM). The solution of lysed nuclei was carefully transferred to 14ml polypropylene centrifuge tubes and centrifuged in an SS-34 rotor at 12,000 rpm for 30 minutes at 4°C. The nuclear extract (supernatant) was removed, its protein concentration determined (described in section 3.2.2.11), and flash frozen in liquid nitrogen. Extracts were stored at –80°C for up to 6 months.

Homogenization buffer:	10mM Hepes (pH 7.6) 25mM KCl 1.8M sucrose ("ultra pure" grade) 1mM EDTA 0.15mM spermine 0.5mM spermadine 0.4mM PMSF 5% glycerol 2µg/ml pepstatin A 2µg/ml leupeptin 2µg/ml aprotonin
0.02 NexB buffer:	20mM Hepes (pH 7.9) 20mM KCl 1mM EGTA 1mM EDTA 5mM DTT 1mM PMSF 10% glycerol 2µg/ml pepstatin A 2µg/ml leupeptin 2µg/ml aprotonin
0.6 NexB buffer	Same as 0.02 NexB buffer, except with 600mM KCl

3.2.2.11 Determination of protein concentration

Determination of protein concentration was carried out spectrophotometrically with protein assay reagent (BioRad). Small volumes (5-10µl) of pure or serially diluted protein preparations were combined with 1000µl assay reagent (diluted 1:5 and filtered). After incubating at ambient temperature for at least 5 minutes, the optical absorption at 595nm (A_{595}) was determined. A standard curve of A_{595} against protein concentration was generated using known concentrations of bovine serum albumin. The standard curve was used to determine the unknown sample concentration.

For very low protein concentrations (less than 1µg/µl), a more concentrated assay reagent was used (1:4 dilution). 5µl of sample was the measured in 100µl volume of assay reagent. These smaller volumes were combined in 96-well plates and the A_{595} values determined with a plate-format spectrophotometer (AbiMed).

3.2.2.12 Generation of tagged hTBP from parent plasmid for western blots

The tagged hTBP transgene product could be generated *in vitro* from its parent plasmid by virtue of the T7 promoter in pAG-17 (see figure 4.2). For this, the TnT coupled reticulocyte system (Promega) was used, providing a convenient size standard (41kDa) for western blotting. The instructions included with the kit were followed exactly to generate tagged hTBP from 1µg of pAG-17 in a 50µl total reaction volume. The reaction mix was aliquoted and stored at –80°C. Generally, 5µl of tagged hTBP reaction mix was used in the size control lane of a western gel.

3.2.2.13 Denaturing polyacrylamide gel electrophoresis (PAGE) of protein

The electrophoretic resolution of proteins was performed under denaturing conditions (0.1% SDS) in a vertical, 1mm thick acrylamide gel slab. The 10x12cm SE280 "mini-gel" system (Pharmacia-Hoefer) provided components for forming and running the gel between two glass plates.

To make a denaturing protein gel, a high-density, low cross-linking "resolving" gel was first prepared (10ml/gel). This was a polyacrylamide solution containing 15% acrylamide, 0.086% bis-acrylamide (achieving a 1:175 cross-linked matrix), 375mM Tris-HCl (pH8.8), and 0.1% SDS. Just before pouring into the cast, 1% (v/v) of a freshly made 10% ammonium persulfate solution and 0.1% (v/v) of TEMED (Sigma) were mixed into the solution. The gel was then poured into the vertical glass plate mold, stopping when the upper level of the solution was 9 cm from the bottom of the gel; about 3 cm from the top of the glass plate. About 800µl of 70% ethanol was gently layered onto the top of this solution, and the gel was left to polymerize. The layer of ethanol ensured an evenly polymerized upper surface. A second "stacking gel" was then prepared, which was a less dense acrylamide matrix with a higher degree of cross-linking. A 5 ml solution (for 2 gels) containing 5.1% acrylamide, 0.14% bis-acrylamide (achieving a 1:36 cross-linked matrix), 125mM Tris-HCl (pH 6.8), and 0.1% SDS. Again, just before pouring the stacking gel, 1% (v/v) of a freshly made 10% ammonium persulfate solution and 0.1% (v/v) of TEMED were mixed into the solution. The ethanol layer was removed from the polymerized resolving gel, and the stacking gel solution was carefully added in its place. The gel cast was filled to the top with stacking gel and the 10 well gel comb was inserted. The gel was allowed to polymerize for at least 2 hours.

To run the protein gel, the comb was carefully removed, leaving formed sample wells in the stacking gel. The bottom surface of these wells was about 1 cm from the interface with the more dense resolving gel. The entire gel and glass plate assembly were clipped into the vertical holder, and 1x TGS buffer was used to fill the anode and cathode chambers. Protein samples were brought to a final volume of 20-30µl with 1x SDS loading buffer. Care was taken to be sure that equal concentrations of salt (KCl) were present in each sample. The samples were then heated at 95°C for 3 minutes before loading into the sample wells with a narrow "sequencing-style" pipette tip. The gel was run at 75-100V until the bromphenol blue had passed through the stacking gel, then the voltage was increased to 150-180V until the dye had reached the bottom of the gel.

10x TGS buffer:	250mM Tris 1.92M glycine 1% SDS
6x SDS loading buffer:	400mM Tris-HCl (pH 6.8) 0.6% SDS 40% glycerol 0.05% bromphenol blue 0.93% DTT

3.2.2.14 Semi-dry electroblotting of protein PAGE onto PVDF membrane

Size resolved proteins in denaturing acrylamide gels were transferred onto PVDF membranes using a semi-dry blotter and a discontinuous buffer system. The gel slab was first removed and immersed in anode buffer II for at least 15 minutes. Dry PVDF membrane was cut in the correct size and first immersed for at least 15 seconds in 100% methanol, then water, and finally for 5 minutes in anode buffer II. Using the anode surface as a support, a block of 6 pieces of electro-blotting paper (Whatman) soaked in anode buffer I was evenly laid down. Onto this was laid 3 more identical pieces of paper soaked in anode buffer I. The PVDF membrane was laid on this, followed by the protein-containing gel slab. 9 pieces of identical size blotting paper soaked in cathode buffer were then evenly laid on top of the gel. The cathode surface was laid on top of the whole transfer sandwich, and the semi-dry blotter was run for 3 hours at 1 mA/cm^2.

Anode buffer I	0.3M Tris 20% methanol (pH not adjusted ~ 10.3)
Anode buffer II	25mM Tris 20% methanol (pH not adjusted ~ 10.3)
Cathode buffer	40mM Amino-n-caproic acid (Sigma) 0.01% SDS 20% methanol (pH not adjusted ~ 7.5)

3.2.2.15 Antibody probing of blotted protein ("western blotting")

PVDF membranes with transferred proteins were blocked overnight in 5% low fat milk in 1xTTBS at ambient temperature on a rocking surface. The membranes were washed the next day at least twice for 5 minutes in 1xTTBS, then probed with the primary antibody in 5% milk/1xTTBS for 2-4 hours under conditions identical to blocking. Membranes were washed at least five times for 5 minutes in 1xTTBS, then treated for 2 hours with diluted peroxidase-labeled secondary antibody under conditions identical to those with primary antibody. The final wash steps were in 1xTTBS for at least five washes, followed by a 10 minute wash in 1xTBS. The membrane was then uniformly covered for 1 minute with a 1:1 mix of the chemilluminescent reagents "A" and "B" (Roche-Boehringer-Mannheim). The membrane was then quickly drained, wrapped in transparent plastic, and exposed to film in a standard light-proof autoradiography cassette.

10x TBS:	0.2M Tris-HCl 1.37M NaCl pH adjusted to 7.6 with HCl
TTBS:	1x TBS with 0.025% Tween-20

3.2.2.16 Silver staining of PAGE resolved proteins

Size-resolved proteins could be visualized in a PAGE slab using silver staining. Gel lanes that were intended for silver staining of proteins generally carried no more than 1µg of total protein, owing to the high sensitivity of the technique. A kit containing all needed reagents for silver staining was used (Pharmacia), and the protocol described in the kit was followed exactly. A stainless steel tray was used for immersing gel slabs in staining solutions. The tray was washed periodically with SDS to exclude any protein contaminants, which can cause background coloring.

3.2.3 Enrichment of tagged hTBP

3.2.3.1 Preparation of P11 ion exchange matrix

P11 ion exchange matrix was produced in lots and stored for future use. 25g of dry P11 resin was added to 2 liters of filtered, de-ionized (milli-Q) water. Gentle swirling followed by gradual settling by gravity at room temperature facilitated swelling and washing of resin. This washing was repeated five times. After decanting away the fifth wash, 2 liters of 0.5M NaOH was added and the resin gently resuspended by slow agitation of the flask. The slurry was allowed to sit at ambient temperature for 5 minutes, then passed through filter paper (Whatman 3mm) using gentle vacuum. The resin was then washed continuously with milli-Q water until the pH of the flow-through decreased to at least 11. The resin was then gently taken up in 0.5M HCL with slow agitation and allowed to sit at ambient temperature for 5 minutes. The slurry was then refiltered through paper as before and washed continuously with milli-Q water until the pH of the flow-through had increased to at least 3. The slurry was then washed on the filter with 1M Tris (pH 7.3)/0.5M KCl until the pH of the flow through was 7.3. The P11 resin was then stored in 0.5M sodium phosphate buffer (pH 7.6) at 4°C for up to three months.

3.2.3.2 Production of 0.9 fraction of nuclear extract

20-25ml of P11 matrix material, which had been prepared earlier (section 3.2.3.1), was gently transferred as a slurry into a 50ml polypropylene tube. The matrix was allowed to settle at 4°C and the aqueous phase was decanted away. Chilled BC300 without protease inhibitors was added (at least three times the settled bed volume) and the tube was gently inverted until an even slurry was formed. The tube was left to settle again at 4°C. This washing was repeated at least 5 times to equilibrate the P11 with BC300.

The protein binding capacity of P11 is normally 15-20mg/ml matrix. Based on this, a required amount of P11 for the intended sample (2-5ml) was carefully transferred to a glass chromatography column (15mm diameter) as a slurry and allowed to settle by gravity (generally 2-5ml). The top of the chromatography column was then covered with an airtight cap with attached 0.8mm silicone rubber inflow tube (BioRad). As a closed system, the matrix bed was further equilibrated overnight at 4°C with fresh BC300 at 2 bed volumes (BV) per hour using a peristaltic pump (Gilson). The next day, the matrix

bed was equilibrated at the same rate for at least two hours with chilled BC300 containing, 1mM PMSF, 2μg/μl pepstatin A, 2μg/μl aprotinin, and 2μg/μl leupeptin. The nuclear extract was then slowly thawed on an ice bath and then pumped through the column at 2 BV/hour. The column was then washed with 5 BV of BC400 without DTT but containing 1mM PMSF, 2μg/μl pepstatin A, 2μg/μl aprotinin, and 2μg/μl leupeptin.

Proteins were eluted at the same pump rate using BC900 without DTT but containing 1mM PMSF, 2μg/μl pepstatin A, 2μg/μl aprotinin, and 2μg/μl leupeptin. 2-3 BV of eluant was collected in series as 0.5-1.0 ml fractions. The concentration of protein was determined for each fraction (described in section 3.2.2.11) and the peak fractions were pooled, flash frozen in liquid nitrogen, and stored for up to 3 months at -80°C.

BC0-BC900 buffers: 20mM Tris-HCl (pH 7.3 at 20°C)
0-900mM KCl
20% glycerol
5mM DTT (added just prior to use)

3.2.3.3 Metal affinity chromatography

Metal affinity chromatography was performed with 0.5-0.75 ml Talon matrix (Clontech). To prepare the column bed, resin was transferred to a glass chromatography column (BioRad) as a slurry and allowed to settle by gravity. The column was closed and a peristaltic pump (Gilson) attached to the silicon inflow tube (0.8mm diameter; BioRad). The Talon matrix was equilibrated at 4°C with at least 5 bed volumes (BV) of BC400 (without DTT) containing 5mM imidazole and 1mM PMSF. The 0.9 fraction was slowly thawed on ice and then appropriately diluted with BC0 (1mM PMSF; without DTT) such that the [KCl] was decreased to 400mM. At 4°C, the diluted extract was then pumped through the column (pump setting 0.8-1.5) three times to ensure thorough binding. The Talon bed was then washed with BC400 (1mM PMSF; no DTT) containing 10mM imidazole for 6-8 BV. The column fraction containing TFIID was then eluted with BC400 (1mM PMSF; without DTT) containing 100mM imidazole. 4-5 BV of eluant was collected in series as 100μl fractions. The concentration of protein was determined for each fraction (see 3.2.2.11) and the peak fractions were pooled. Aliquots were made as needed, to which DTT and acetylated BSA (Sigma) were added to 5mM and 100ng/μl, respectively. The aliquots were flash-frozen in liquid nitrogen and stored at -80°C.

3.2.3.4 *In vitro* transcription assays

In vitro transcription reactions were carried out essentially as described previously (110;122;123) with purified transcription components from numerous sources. The α and β subunits of TFIIE as well as the TFIIB monomer were produced from human cDNAs, recombinantly expressed in *E. coli* and purified through introduced histidine-stretch tags (122). TFIIF and TFIIH were co-purified from Hela cell nuclear extract through multiple chromatographic steps outlined elsewhere (122;124).

Naturally occurring TFIID from Hela cell nuclei was needed for control reactions and was purified as described elsewhere (122). RNA Pol II holoenzyme was purified from new-born calf thymus essentially as described by Kim and Dahmus (125) with some modifications (123). Activated *in vitro* transcription with human components depends of the presence of positive co-factors (PCs). The factor PC4 used in this work was purified from recombinant expression of a human cDNA in *E. coli*, as described elsewhere (126). GAL4/Sp1 is a fusion protein consisting of the N-terminal 94 amino acids of yeast GAL4 protein and the complete activation domain of human Sp1. This hybrid protein has been used previously for PC-mediated activation of *in vitro* transcription studies on promoters containing GAL4 binding sites (127).

An alternative to the use of highly purified components was the use of a "0.5 system fraction" for some *in vitro* transcription reactions. This consisted of the 0.3-0.5M KCL elution from P11 fractionation of Hela nuclear extracts (110;123). This fraction contained all but two of the GTFs needed for basal transcription; TFIIE, TFIIF, TFIIH, and RNA Pol II. It also contained PC2 (122), a positive cofactor that, like PC4, is able to potentiate Sp1-activated transcription (G. Stelzer, personal communication). Thus, this "system fraction" was useful to test for TFIID activity in affinity-purified hTBP preparations, requiring only purified GAL4/Sp1 and TFIIB to complement the reaction.

The plasmid pMRG5 contained the DNA template that was used for activated transcription reactions (122). It consists of 5 GAL4 recognition sites (128) upstream of the HIV-1 TATA region. This promoter is linked to a downstream 400 base pair "G-less" reporter gene, containing only the three bases C, A and T on the sense strand (129). This artificial gene is designed for transcription reactions supplemented with only UTP, CTP and ATP. At the end of the G-less region the polymerase is halted due to lack of GTP, and mRNA of a fixed length (~400 bases) accumulates. To monitor changes in core promoter activity, *in vitro* transcriptions also contained the AdML core promoter lacking recognition sites for regulatory factors. This resided on a separate plasmid (pMLΔ53) where it directed expression of a truncated version of the G-less reporter gene (123). All necessary transcription factors and DNA templates used in this study were generous gifts of G. Stelzer and M. Meisterernst.

Transcription reactions carried out *in vitro* with purified components contained 20ng TFIIE (10ng of α and β subunit), 20ng TFIIB, 150ng TFIIF/TFIIH (co-fractionated), 10-20ng RNA pol II, 50ng PC4 and ~50ng GAL4/Sp1. 10ng Talon purified hTBP/TFIID (see 3.2.3.3) was used for each reaction, or, where appropriate, 100ng of TFIID purified from Hela nuclei was used as a control. The addition of TFIIE, TFIIF, TFIIH, RNA Pol II, and PC4 could be omitted if 4μg of Hela "0.5 system fraction" (described above) was used.

Generally, each *in vitro* transcription reaction was begun with 40ng of each DNA template in a buffer containing 5mM $MgCl_2$, 60mM KCl, 25mM Hepes/KOH (pH 8.2), 10% glycerol, 0.2mM PMSF, 5μM DTT, 0,1μg/μl acetylated BSA (Sigma) and 20 units/reaction of RNAsin (Roche-Boehringer-Mannheim). As necessary, activator (GAL4/Sp1) and cofactor (PC4) were added to this. A "system mix" was separately prepared for each reaction which contained the needed GTFs and "0.5 system

fraction" as required. The system mix had the same final buffer components and concentrations as the DNA mix and was combined with the separate DNA-containing reaction tubes to give a final volume of 20µl. After 30 minutes incubation at 28°C, 1µl of the ribonucleotide mix was added to start each reaction, which were allowed to run for 60 minutes at 28°C. The reactions were stopped with the addition of 400µl stop buffer, extracted once with phenol-chloroform, and precipitated for at least 1 hour at –20°C with isopropanol. The transcription mix was centrifuged at 11,000rpm for 30 minutes, the pellet washed once with ice cold 70% ethanol and then vacuum dried. 10µl of transcription loading buffer was added and the pellet was resuspended after ~15 minutes at 60°C.

Ribonucleotide mix:	0.2µl OmG-mix 0.8µl α[^{32}P]CTP (3000Ci/mmol) (Amersham)
OmG-mix:	10mM ATP (Amersham) 10mM UTP (Amersham) 0.5mM CTP (Amersham) 2mM 3'-O-methylguanosine-5'-triphosphate (Pharmacia)
Stop buffer:	10mM Tris (pH 7.8) 10mM EDTA 0.5% SDS 100mM LiCl 100µg/ml *E. coli* tRNA (Roche-Boehringer-Mannheim) 300mM sodium acetate
Loading buffer:	97% deionized formamide 20mM Tris-HCl (pH 7.0) 10mM EDTA 0.03% bromphenol blue 0.03% xylene cyanol

3.2.3.5 Gel electrophoresis of *in vitro* transcribed mRNA

^{32}P-labeled mRNA probes that were generated from *in vitro* transcription were visualized by denaturing polyacrylamide gel electrophoresis (PAGE). Using a standard 20x16cm vertical gel electrophoresis assembly, a gel matrix was formed which contained 5% polyacrylamide, 7M urea and 1x TBE. The steps prior to pouring the gel into the glass plate assembly are largely identical to those for PAGE of S1 nuclease protected probes (see 3.2.2.9). The gel was allowed to polymerize for at least two hours using a comb to form flat-bottom sample wells. The gel was then pre-run for 30 minutes at 50mA. Samples were then loaded and run on the gel for 90 minutes at 50mA. The gel slab was then removed from the glass plates, fixed for 15 minutes in 15% acetic acid, then washed for 5 minutes in distilled water. The gel was then affixed to 3MM-type chromatography paper (Whatman) and dried under vacuum at 70°C for 1 hour. The bands of labeled RNA in the dried gel were then visualized with either standard film radiography or a Phosphor-imager (BioRad).

4. RESULTS

4.1 Creation of transgenic mice

In order to create the transgenic animal model as described in section 2.14, we obtained a "tagged" cDNA construct of human TBP (a kind gift of A. Goppelt and M. Meisterernst; Universität München). This was the full length cDNA for hTBP (56) with 96 base pairs (32 codons) added to the 5' end. The translated 5' sequence included an AUG start site as well as two affinity purification tags: the "HA" epitope sequence and a stretch of 6 histidine residues (see figure 4.1).

The HA sequence is an antigenic epitope derived from the hemagglutinin protein of the human influenza virus (130). Antibodies against this epitope are commercially available (Roche-Boehringer-Mannheim) and have been used routinely for immuno-purification of HA-tagged proteins (131), including TBP from cell lines (71;132) and *Drosophila* embryos (133).

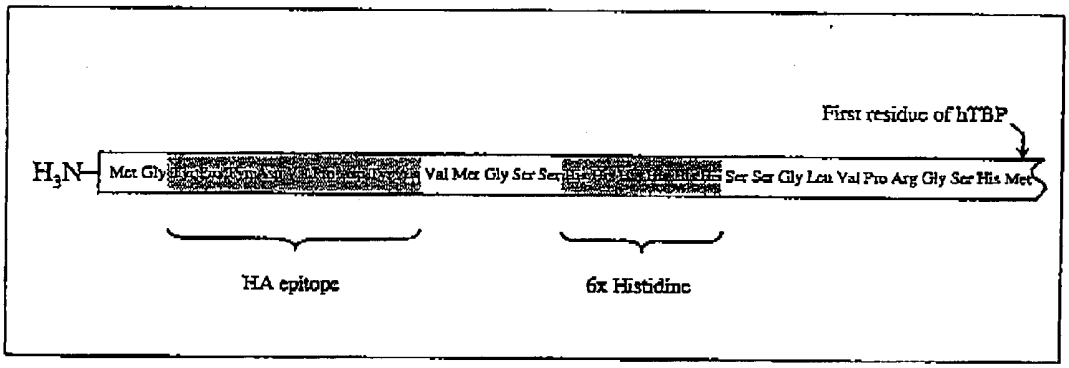


Figure 4.1: Diagram of N-terminal double tag region of transgenic hTBP protein.

The six consecutive histidine residues also function as an affinity purification tag based on interactions between some electropositive transition metals (including Co^{2+}, Ni^{2+}, Zn^{2+}, and Cu^{2+}) and histidine. Such metals can be affixed to porous resins to take advantage of the reversible binding of 6 consecutive histidine residues, also known as immobilized metal affinity chromatography (IMAC) (134-136). Various IMAC resins for such purification are available commercially, including the Talon affinity resin from Clontech which was used in this work. Histidine-Talon interactions can be both formed and reversed while maintaining native conditions by addition (or exclusion) of 50-100mM imidazole.

The decision to use a double-tagged TBP transgenic protein was based on the intent of eventually generating a very pure TFIID preparation. As mentioned above, HA-tagged TBP has been successfully used to recover TFIID from cell culture expression. However, it was reasoned that in a whole animal model, the availability of alternate purification strategies may be indispensable to recovery of holo-TFIID from an array of different whole tissue. Further, it was assumed that the recovery of transcriptionally

competent TFIID complex from whole tissue could be enhanced by multiple affinity chromatography steps, each carried out with relatively mild washing conditions.

4.1.1 Design of transgenic models

Until now, a whole-animal model which transgenically expresses extra versions a central component of the transcription complex (such as TBP) has not been created. It was speculated that a stoichiometric disturbance at the level of GTFs would be potentially deleterious, resulting in non-viable embryos or unhealthy mice. To avoid such pitfalls, it was decided that two separate transgenic lines should be attempted; one in which tagged hTBP expression can be induced shortly prior to tissue collection, and one in which the transgene is expressed constantly. The correct expression of any such transgene requires the coding sequence (usually a cDNA) to be bracketed on the 5' end by an appropriate promoter, and on the 3' end by a polyadenylation (polyA) site. Plasmids containing numerous combinations of promoter and polyA sequences for either inducible or constitutive transgenic expression are widely available.

4.1.1.1 Construction of inducible-expression hTBP transgene

For inducible expression of tagged hTBP, the promoter for the mouse metallothionine I (MT-1) promoter was used. This promoter is generally inactive, but will induce expression in an adjacent gene in the presence of metals, including zinc and cadmium (137), (138). This property makes it an attractive choice for transgenic models, where it can be induced to express a transgene at high levels shortly after injection of the animal with $ZnSO_4$or $CdSO_4$ solutions (139-143).

To create an MT-1 promoter/tagged hTBP construct, we used the commercially available pBPV plasmid (Pharmacia; GenBank accession #U13843). This promoter contains the MT-1 promoter as well as the Moloney sarcoma virus (MSV) long terminal repeat sequence, a strong mammalian enhancer element (figure 4.2). In addition, there is a downstream polyadenylation sequence from the simian virus 40 (SV40) early locus. Between the promoter and the polyadenylation sequences lies a multiple cloning sequence, which contains several restriction sites for convenient ligation of cDNA. The tagged hTBP was excised from its parent plasmid pAG-17 with NcoI and NotI, yielding a 1325 base pair fragment. The "vector" pBPV plasmid was digested with XhoI and NotI. The linear vector and insert DNA strands were gel purified and ligated together. Because the restriction sites in the MCS did not match both ends of the insert, an oligonucleotide "adaptor" was synthesized and hybridized to link the 5' end of the tagged hTBP with its vector host, as indicated in figure 4.2.

To excise the final DNA construct for microinjection, the pBPV-hTBP construct was digested with ClaI and BamHI, yielding the 3249 base pair fragment containing promoter, cDNA and polyA sequence (figure 4.2, bottom). The DNA fragment was isolated by native gel electrophoresis and extensively purified through phenol-chloroform extraction prior to microinjection of fertilized mouse oocytes.

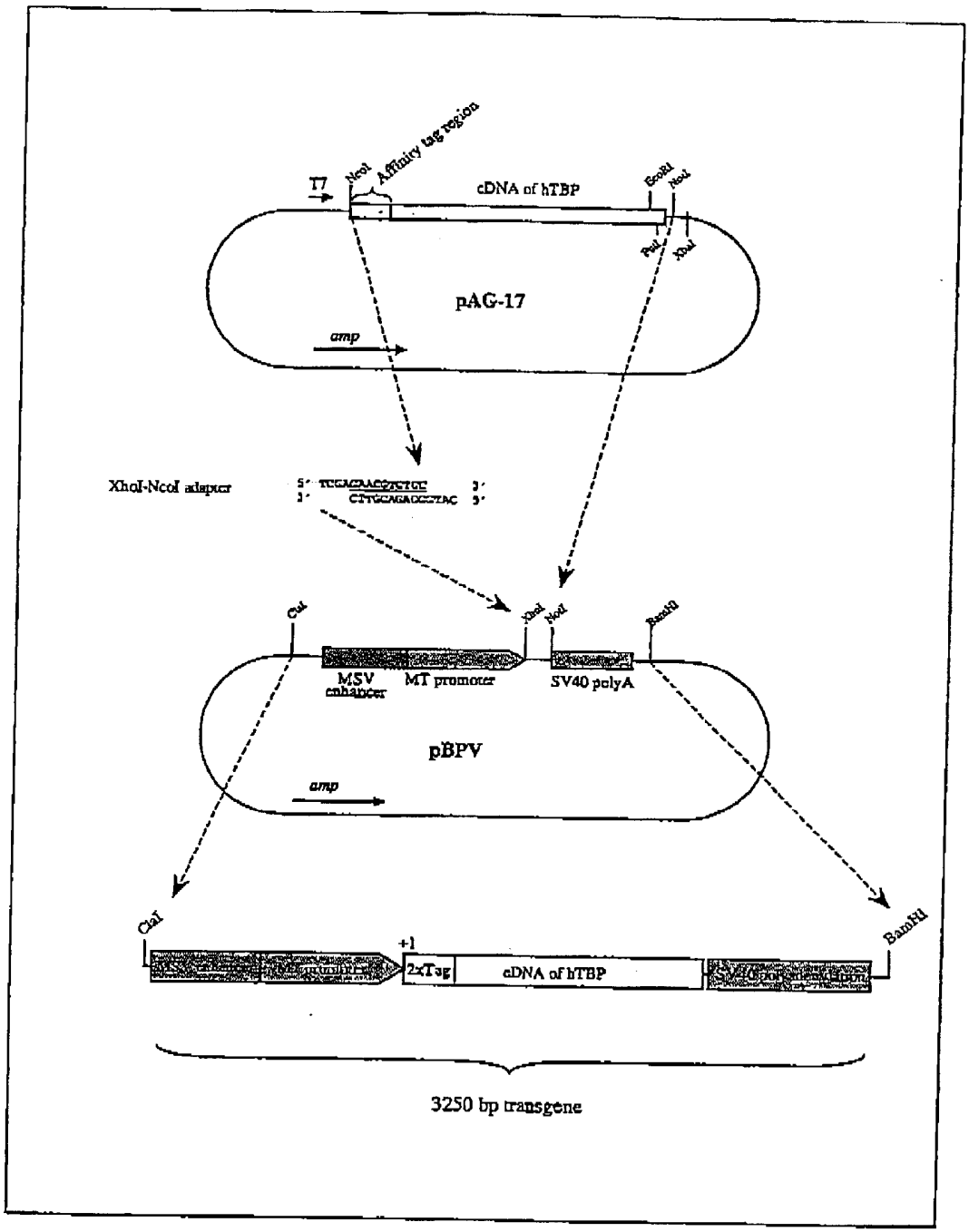


Figure 4.2: Schematic representation of construction of MT-hTBP transgene. Genetic elements are not drawn to scale.

4.1.1.2 Construction of constitutive-expressing hTBP transgene

To express the tagged hTBP cDNA in a constitutive manner, the promoter for the human Elongation Factor-1-alpha (EF-1α) locus was chosen (144). This is a promoter that has been shown to strongly express in most mammalian tissue at a constant level (145-147). The plasmid pEFBG (kind gift of H. Greg Polites; Hoechst Marion Roussel, USA) contains the EF-1α promoter and a downstream polyadenylation sequence from the bovine growth hormone (BGH) gene (see figure 4.3). These two elements are separated by a multiple cloning sequence containing NcoI and NotI sites. Since the final excision enzyme was intended to be EcoRI, it was necessary to first use site-directed mutagenesis to abolish the hTBP-internal EcoRI site. The requirements for this base change protocol require the mutagenesis of a unique restriction site in the plasmid (see 3.2.1.7). Therefore, both the 3' EcoRI and NotI sites were abolished in pAG-17, giving pAG-17*mut*. The tagged hTBP cDNA could them be removed from the parent plasmid as a NcoI-XbaI fragment and gel purified. Likewise, the pEFBG vector plasmid was digested with NcoI and NotI and gel purified. The ligation of the NcoI-XbaI fragment into the vector was facilitated by the use of a synthesized oligonucleotide XbaI-NotI adapter, as indicated in figure 4.3.

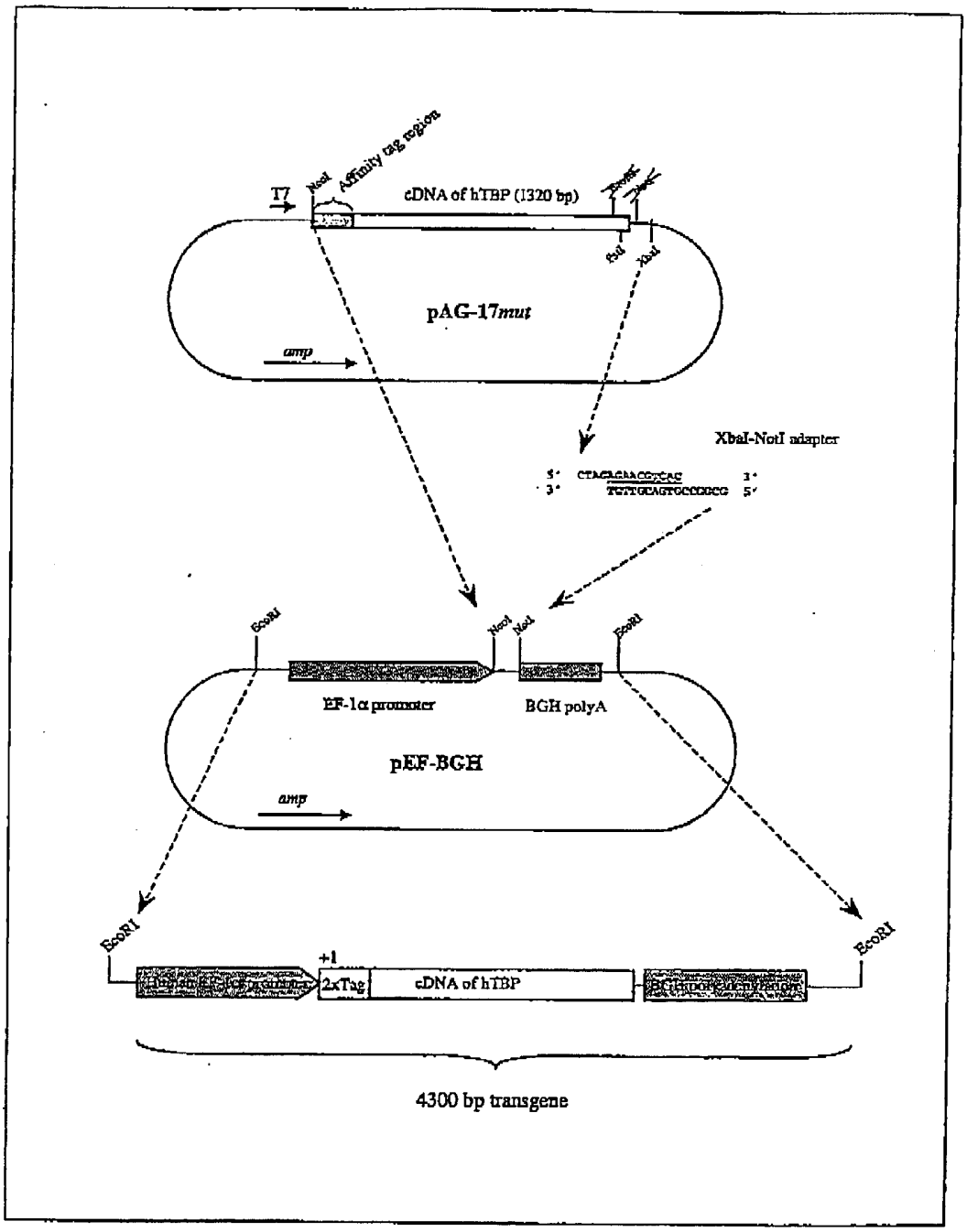


Figure 4.3: Schematic representation of construction of EF-hTBP transgene. Genetic elements are not drawn to scale.

To excise the final DNA construct for microinjection, the pEFBG-hTBP plasmid was digested with EcoRI, yielding a 4300 base pair fragment which was isolated by native gel electrophoresis (figure 4.3, bottom). The DNA was purified by extensive phenol-chloroform extraction prior to microinjection.

4.1.2 Determination of transgene-containing founder mice

Microinjected mouse zygotes were carried to term in surrogate mothers (see 3.2.2.1) (121). These first generation (G_0) mice were all retained and genomic DNA collected from tail tissue between 2 and 4 weeks of age (see 3.2.2). PCR was carried out on the genomic DNA (see 3.2.2.3) to determine which G_0 mice carried the microinjected DNA. For the MT promoter construct, PCR primers annealing at positions -150 and +430 were used, yielding a 580 bp product for transgenic G_0 mice. For the EF-1α promoter construct, primers annealing at -70 and +430 were used, yielding a 500 bp product for positive G_0 mice. PCR products were visualized with gel electrophoresis as described in section 3.2.1.13.

Of 150 possible G_0 mice with the MT/TBP transgene, 36 (24%) were carrying the gene (data not shown). For the EF-1a/TBP mice, 63 possible G_0 mice were tested. Of these, 9 (14.3%) were positive (data not shown).

4.1.3 Detection of transgenic mRNA with S1 nuclease protection assay

The integration of an untargeted DNA construct into the host genome is an inherently random event. A common occurrence in the establishment of a transgenic animal line is the insertion of the transgene into the genome in such a way that it is rendered inoperative. In order to determine if the G_0 founder mice were also able to generate tagged hTBP transcript, total RNA was assayed for the presence of transgenic mRNA. G_0 mice were retained until mature and then outbred with non-transgenic mice. Resulting litters were therefore ~50% transgenic and needed to be tested for presence of the transgene using PCR, as described above. Some of the positive, heterozygous G_1 mice corresponding to each positive G_0 founder mouse were sacrificed and total RNA isolated from liver, brain, spleen, heart, kidney, and muscle. Transgenic mice carrying the MT promoter were induced four hours prior to sacrifice, as described in section 3.2.2.4.

The use of standard northern RNA analysis using longer (>200 bases) DNA probes would likely present difficulties, given the strong sequence similarity between transgenic (human) and endogenous (mouse) TBP. Therefore, the presence of transgenic mRNA was detected in liver RNA preparations by S1 protection assays, a method which is sensitive to shorter target sequences (50-70 bases). The 76 base oligonucleotide S1/5' was used (see section 3.1.11) which hybridized to part of the unique 5' tag region of the hTBP transcript. The first 12 bases on the 3' end of S1/5' were specifically designed to be non-complementary to transgenic mRNA. Such a non-matching, single stranded "tail" would be susceptible to S1 nuclease digestion, even when the primer has specifically hybridized with the mRNA target. This provided an internal control which ensured that the S1 nuclease has indeed digested all single-stranded nucleic acid, thus eliminating the possibility of false positives (see figure 4.4 B). RNA samples containing the tagged hTBP transcript yielded an S1 nuclease protected band of 64 bases. Conversely, the failure to express tagged hTBP mRNA was recognized by the absence of a protected band. Analysis of liver RNA identified 6 (out of 36) transgenic lines which transcribe the MT promoter construct (MT-hTBP), and 4

(out of 9) lines which transcribe the EF-1α promoter construct (EF-hTBP). Typical examples of this screening are shown in figure 4.4 A.

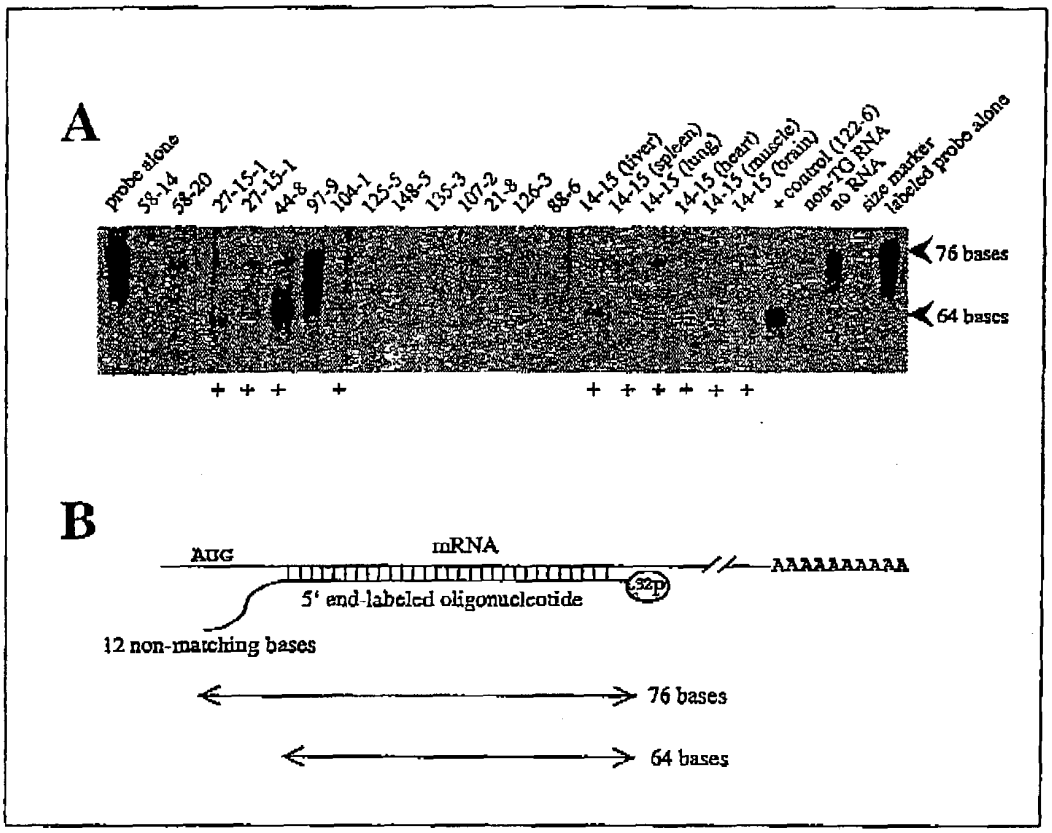


Figure 4.4: Typical results of some S1 nuclease protection assays using the S1/5' oligonucleotide. A schematic representation of the assay is depicted below.

An obstacle to creation of transgenic models is the occurrence of "mosaic" animals which retain or express the transgene in only a subset of tissues. This can be caused by numerous unpredictable integration-related phenomena. To guard against this possibility, S1 nuclease protection assays using probe S1-5' were used to probe RNA from other tissues of G_1 mice. In all cases, transgenic mRNA was present in liver, spleen, lung, heart, muscle, brain. Typical results for one of the MT-hTBP lines is shown in on the right side of figure 4.4 B.

The production of incomplete transcripts, through cryptic termination sites or other unpredictable gene integration effects, is another common pitfall in the establishment of a transgenic model. To rule this possibility out, further S1 nuclease protection assays were conducted using primer S1/3' (see section 3.1.11). This oligonucleotide hybridizes with a unique region of the 3' transcript, downstream of the translation "stop" codon. Analogous to S1/5', the first 12 bases of the 3' end of S1/3' were designed to be

non-complementary to transgenic transcript, yielding a 62 base long protected oligonucleotide when the targeted mRNA was present. All of the animals that were positive with the S1/5' primer were also positive with the S1/3' primer. This indicated that all lines being studied were producing full-length transgenic mRNA. Typical examples of these assays are shown in figure 4.5.

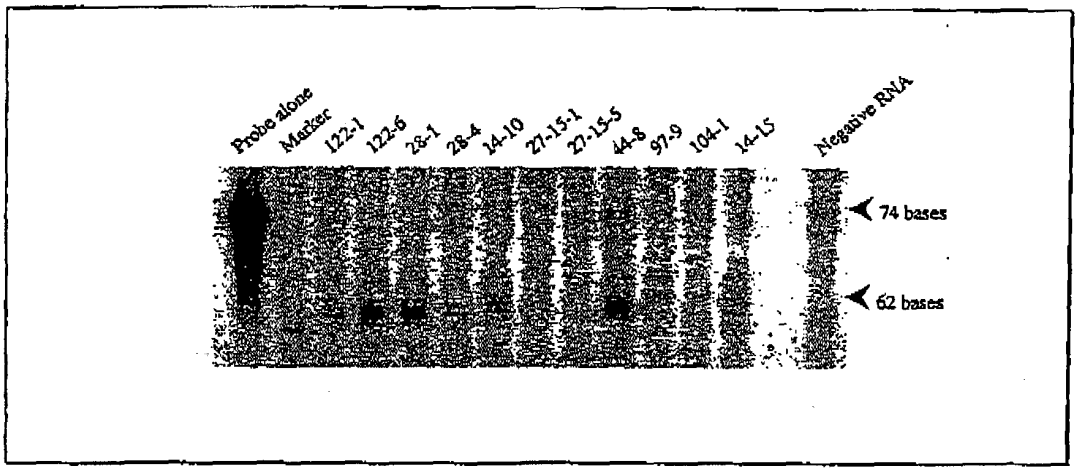


Figure 4.5: Typical transgenic screening results of S1 nuclease protection assay using S1/3' oligonucleotide.

4.1.4 Detection of hTBP transgene product

The six MT-hTBP and four EF-hTBP mouse lines which proved able to create hTBP transcript were outbred to further generations using non-transgenic mates. All litters were therefore ~50% transgenic, and consequently needed to be PCR-tested for the inheritance of transgene (data not shown). Groups of mice from G_2-G_4 litters which were determined by PCR to be carrying the transgene were then sacrificed and nuclear extract from individual livers was prepared. MT-hTBP mice were induced to activate the transgenic promoter prior to sacrifice. The extracts were tested with western blotting for the presence of hTBP using anti-HA antibody as the primary antibody. Western analysis of many litters was carried out to determine which transgenic line(s) had optimal expression of transgene for larger-scale purification of TFIID. Tagged hTBP was transcribed and translated from the T7 promoter of the pAG-17 plasmid, providing a convenient protein size marker. A typical western blot of some MT-hTBP and EF-hTBP liver nuclear extract is shown in figure 4.6.

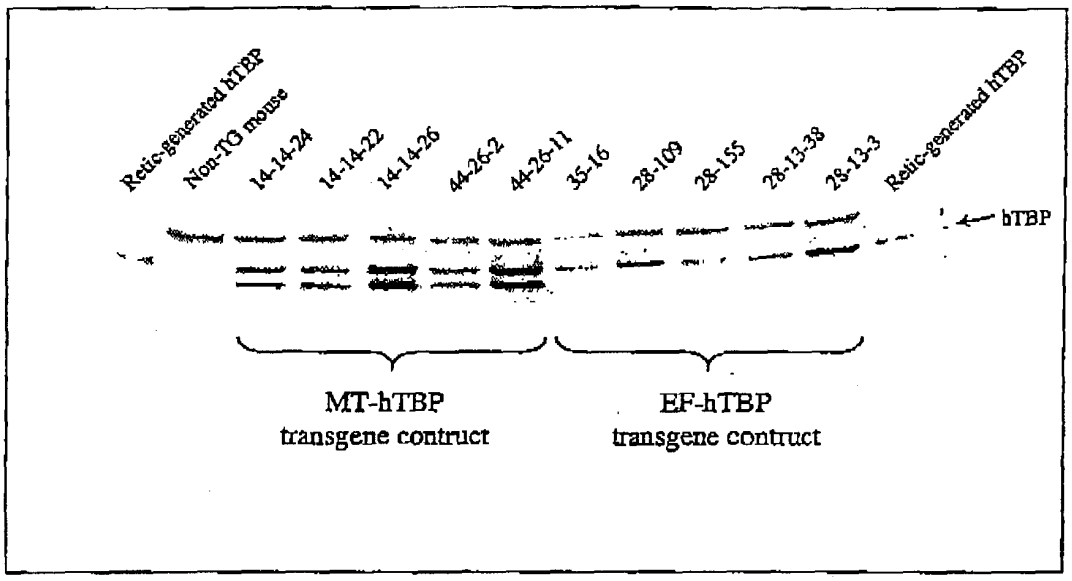


Figure 4.6: Typical western blot detection of tagged hTBP in nuclear extracts from transgenic liver.

In total, offspring from 4 of the 6 mRNA-containing MT-hTBP lines consistently contained tagged hTBP protein (data not shown). A lower molecular weight version of tagged hTBP appeared as a second band in all four the protein-containing MT-hTBP mouse lines, most likely indicating a truncated variant. Since it appeared consistently in all four lines, it was thought to be an effect that is intrinsically related to the MT-hTBP construct, and not a random integration-related effect.

In the case of the EF-hTBP lines, all four mRNA-containing lines produced offspring which contained the transgene product. However, offspring from three of them consistently expressed comparatively lower levels of the protein than the fourth line (data not shown).

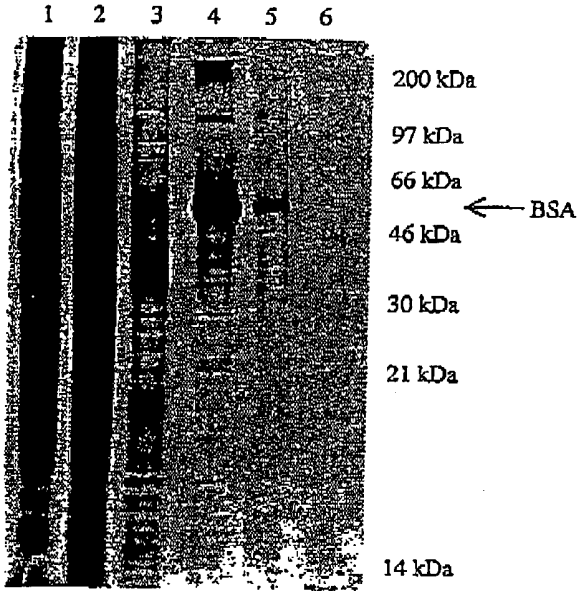
Western blotting was also used to test nuclear extracts in brain, kidney, and lung of all 4 MT-hTBP lines and the strongest expressing EF-hTBP line (line #28). Correctly sized transgenic protein was present in these tissues as well (data not shown).

4.1.5 No obvious ill effects of transgenic hTBP expression

As mentioned in 4.1.1, a projected pitfall of over-expressing TBP in a transgenic mouse was the uncertainty of the stability and health of the animals. The far-reaching implications of disturbing the natural levels of TBP in a whole animal can not be comprehensively predicted at the outset of such a study. The failure of this transgenic model due to developmental defects or cancers would not have been unexpected. Transgenic tissues were not subjected to formal histological evaluation. However, no obvious defects or health problems were noted in the ~300 mice that were used, many of which reached 18 months of age before being sacrificed.

4.2 Enrichment of tagged hTBP

The apparent health of the transgenic mice led to the acceptance of the constitutive-expressing tagged hTBP mouse line, EF-hTBP line #28, as the favored model for further TFIID purification efforts. Heterozygote EF-hTBP mice between 4 and 10 months old were sacrificed, and nuclear extracts from pooled liver, brain, and kidney were prepared. Liver from 20-30 animals was used (30-35g tissue), giving a final yield of 200µg of hTBP-enriched fraction. Brain and kidney from 60-100 animals was used (20-25g), yielding 100-150µg of hTBP enriched fraction. The finished nuclear extracts were then enriched for TFIID, using P11 chromatography to generate a 0.9M KCL fraction (see section 3.2.3.2). These fractions were then subjected to Talon affinity chromatography under native conditions in order to enrich holo-TFIID associated with the tagged hTBP. The fractionation of crude nuclear extract could be visualized by silver-staining of SDS-PAGE resolved total protein in each fraction (see figure 4.7).



Lane 1: 1µg crude liver nuclear extract
Lane 2: 1µg P11 loading flow-through
Lane 3: 1µg BC900 pooled peak fractions
Lane 4: 1µg Talon loading flow-through
Lane 5: 1µg Talon pooled peak fractions

Figure 4.7: Silver stained SDS-PAGE showing enrichment of fractionation of total protein content in liver nuclear extract from EF-hTBP transgenic mice. The thick protein bands between 46 kDa and 66 kDa in lanes 4 and 5 are from BSA, which was added to the samples to stabilize protein.

The successful enrichment of transgenic protein could be monitored by subjecting comparable amounts of total protein from the various chromatographic steps to western analysis (see figure 4.8).

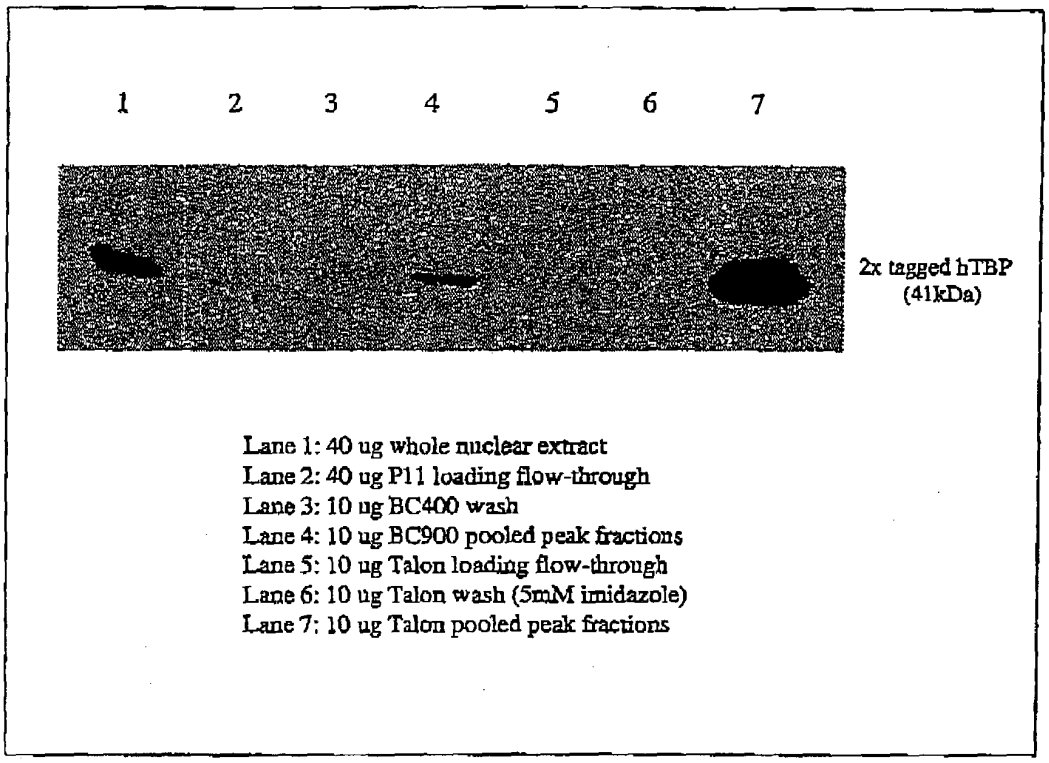
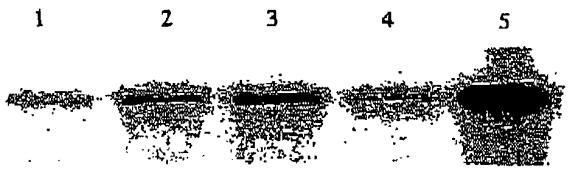


Figure 4.8: Western blot showing enrichment of hTBP through fractionation and purification steps of transgenic liver nuclear extract. High-affinity anti-HA rat mAb was used for detection.

4.3 Transgenic hTBP/TFIID from different tissues is transcriptionally competent

To determine if native holo-TFIID can indeed be enriched through the tagged hTBP subunit, Talon chromatography enriched fractions (see section 4.2) were tested for TFIID activity. Standard *in vitro* transcription reactions were performed, replacing human TFIID with the Talon-enriched fraction from liver. Combinations of highly purified transcription components were used to monitor both basal and activated transcription. Figure 4.9 illustrates the ability of liver tagged hTBP/TFIID to support basal (lane 2) as well as activated transcription (lane 5). Some background activity is normally present in such *in vitro* systems (lane 1) and is attributed to minute amounts of human TFIID that often contaminates GTF preparations derived from Hela cells. Identically prepared Talon fractions from non-transgenic mouse liver showed negligible activity compared to those of the EF-hTBP mouse line.



Lane 1: Background activity of GTFs without TFIID
Lane 2: Basal transcription with 6xHis purified TFIID
Lane 3: TFIID activity in presence of GAL4/Sp-1 activator
Lane 4: TFIID activity in presence of PC4 activator
Lane 5: TFIID activity in presence of GAL4/Sp-1 and PC4

Figure 4.9: Basal and activated *in vitro* transcription using Hela GTFs and TFIID derived from transgenic tagged hTBP from liver.

Similar transcription assays were used to determine if holo-TFIID from other organs would also be transcriptionally competent. Transgenic hTBP/TFIID was enriched as described above (section 4.2) from liver, brain and kidney. Transcription assays were performed with these extracts and a "0.5 system fraction" from Hela nuclei (described in section 3.2.3.4). As with liver, preparations from brain and kidney also supported both basal and activated transcription. A normalized profile of hTBP/TFIID activity between organs was made by comparing the radiolabeled mRNA of activated transcription to that of basal transcription. The results of such a comparison are shown in figure 4.10. The bar on the far right shows the typical increase in transcription with the human components upon activation with GAL4/Sp1induction. (~3 fold). Liver and kidney showed similar increases while brain had the strongest activation response (~7 fold).

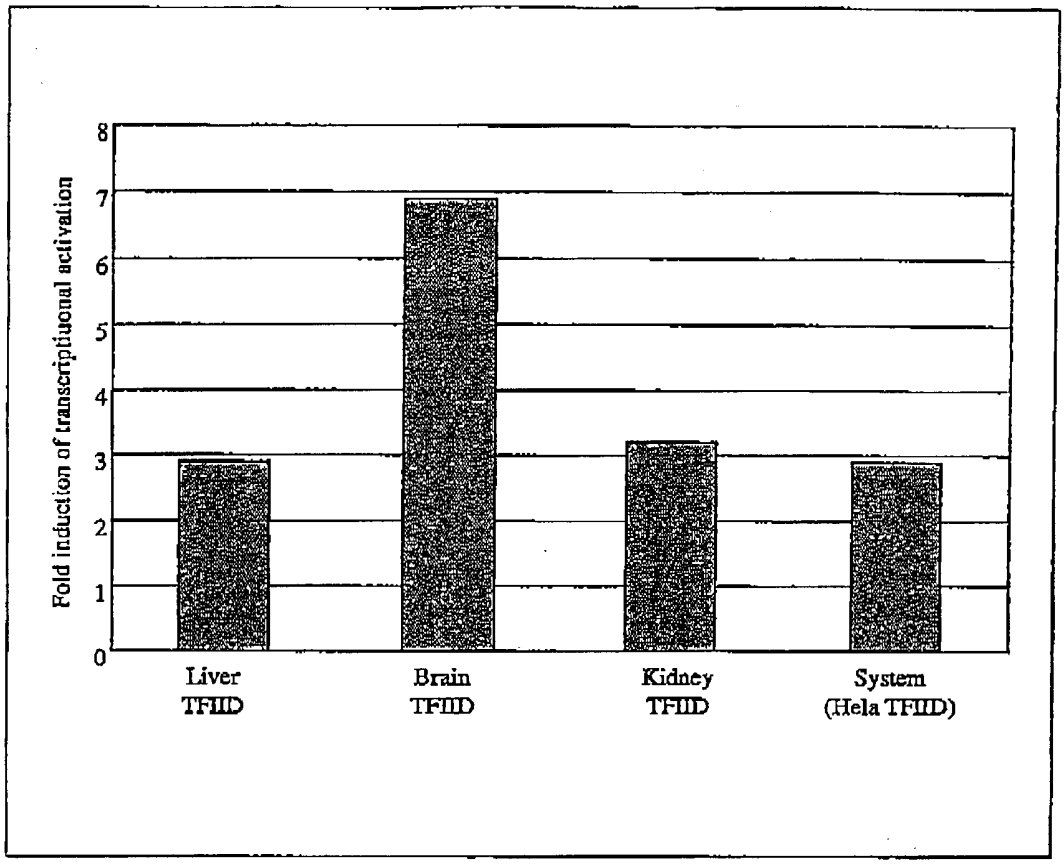


Figure 4.10: The abilities of various affinity-derived TFIIDs to support activated transcription are compared to their abilities to support basal transcription (fold induction).

5. Discussion

5.1 Conclusions

In this thesis work, initial steps have been taken to develop a novel approach to studies of eukaryotic mechanisms of RNA transcription. A model has been described in which affinity-tagged TBP is transgenically expressed in a whole organism, an approach that extends from analogous studies in cell lines (71;132;133).

The results outlined in sections 4.2 and 4.3, taken together, show that there is a TFIID activity which is co-purified with the transgenic tagged hTBP protein. This TFIID supports archetypal regulated transcription, namely activator-mediated activation of transcription. As discussed in section 2.6 and 2.7, the ability of such regulatory factors to activate transcription is dependant on the presence of certain TAF_{II} subunits in the holo-TFIID. It has been reported that the TBP subunit alone can not support activated transcription (Martinez et al., 1994). A comparable TFIID activity is not present in nuclear extracts from organs of non-transgenic mice that have been subjected to parallel fractionation and chromatography. Hence, the demonstrated TFIID activity is attributed to endogenous mouse TAF_{II}s associating with the transgenic hTBP. This novel system is successful in that it enables a transcriptionally competent holo-TFIID to be efficiently obtained from different vertebrate tissues. Further, that this model permits the possibility of convenient and direct access to TFIID-associated proteins from multiple vertebrate tissues, some of which may be found to be tissue-specific factors.

Some of the specific attributes of this system allow observations which should prove useful in pursuing this transgenic model, or related ones, in the future. Perhaps the most immediate useful comment that this system provides is the fact that a transgenic TBP mouse is indeed possible. As mentioned earlier, the failure of this model due to unpredictable negative effects of TBP over-expression would not have been unexpected. Since the mice appear normal and healthy, as noted in section 4.1.5, constitutively-expressing tagged TBP systems in other species may hereafter be designed with more certainty. Thus, the transgenic mouse line described in this thesis can be considered a "proof of concept" for further related studies.

As mentioned in section 2.6, mammalian TBP protein has highly conserved homologs in evolutionarily distinct organisms such as yeast and *Drosophila*. Indeed, numerous reports indicate that many of the subunits retain functional conservation between diverse species (see figure 2.4). For example, functional holo-TFIID can be assembled from yeast TBP and human TAF_{II}s (76). In this study, it is demonstrated that a human cDNA for TBP can be expressed and is functional in a different mammalian (mouse) nucleus. This is perhaps not surprising, given the particularly well-conserved sequence similarity between mouse and human TBP (59). However, a long-term goal of using a human transgene in this non-human, mammalian system was to establish a paradigm for subsequent related studies. Future applications for this model have been anticipated (discussed below) in which a human TBP is expressed in other vertebrate species in an effort to efficiently purify large amounts of holo-

TFIID. For such studies, the use of an established "standard" tagged TBP would be useful for comparison of observations from different systems. Human TBP is a sound choice for such a standard since it has been established that an N-tagged version will fold correctly, attract available TAF_{II}s, and function as TFIID in cell culture (71;132;133). With the model developed in this thesis work, this certainty now extends to a whole animal system. Furthermore, some of the future applications for this system may be related to pharmaceutical research (discussed below). Therefore, the use of human components in such a system, whenever possible, is advantageous. Since the tagged hTBP has now been shown to work in mice, there will be less uncertainty associated with the creation of analogous transgenic models in other larger species.

The affinity-tag purification described in this work differs from previous tagged hTBP studies in that a "six histidine stretch" tag has been used. Metal affinity chromatography (Talon matrix) enables efficient and rapid enrichment of the tagged protein from P11-fractionated nuclear extract (see section 4.2). It should be noted that the eluted TFIID components are not highly purified after this step. Total protein analysis (SDS-PAGE, silver staining) showed an array of contaminating proteins in identically treated extracts from transgenic and non-transgenic mice. Attempts were made to purify the TFIID further through the HA epitope tag using commercially available anti-HA resin (Babco) and established HA-tagged hTBP purification techniques (71). Indeed, available transgenic hTBP was able to bind to the antibody matrix, but detectable amounts could be eluted only under denaturing conditions. The ability to elute holo-TFIID from the resin by epitope- peptide competition was hampered, presumably by the high-affinity of the antibody and limiting amounts of starting material. Previous native TFIID studies using this affinity procedure have had the advantage of using Hela cells, which can be grown in large volumes to prepare nuclear extracts. Most of these studies used nuclear fractionates from >20 liters of high density cultured Hela cells prior to affinity purification of TFIID by way of an epitope tag. It is inferred that only with such comparatively large amounts of source material is it suitable to use epitope peptide competition to elute native, tagged protein from the antibody matrix. However, it is understood that the HA tag will become more useful for purification when the mouse model is available as a homozygotic line (currently being developed) and, therefore, more tissue can be reasonably generated for TFIID preparation. Similarly, the HA tag will prove instrumental in the use of this transgene in an animal with more abundant source tissue, such as rat or swine.

5.2 Future prospects

5.2.1 RNA Polymerase II studies

As discussed earlier, the majority of transcription studies have focused on components and mechanisms in the *Drosophila* embryo, the human Hela cell line, and yeast. Models based on observations from each of these experimental systems have individual advantages and disadvantages (which have been the source of animated debates between research groups). Taking the results of

biochemical and genetic studies from all three systems together, however, an intriguing and detailed picture of fundamental transcriptional events has emerged over the last 20 years. The transgenic system described in this thesis work represents a first step toward a new and more comprehensive approach for the investigation of the underlying principles of regulated transcription in higher eukaryotes. This step is a significant departure from the traditional unicellular experimental systems (Hela cells, *Drosophila* embryos and yeast). The demonstrated ability to efficiently purify transgenic TBP as a functional TFIID activity indicates the feasibility of pursuing such investigation in the context of a metazoan system.

Properly regulated transcription requires, among other things, the complex network of activities and interactions associated with the TAF_{II} subunits of TFIID. These factors confer promoter preference and other gene-specific effects on the TFIID complex. Expression of some TAF_{II}s has been observed to be restricted to certain tissues (see section 2.13), thus providing a plausible mechanism for tissue-specific gene regulation. Preparations of TFIID from a transgenic model as described in this work will become a powerful tool to confirm and extend current regulatory models to include metazoan considerations. Through the use of both affinity tags, highly purified holo-TFIID from different tissues can be qualitatively evaluated with denaturing gel electrophoresis. Isolated TAF_{II} bands could then be subjected to micro-sequence analysis and compared to those TAF_{II}s which have been previously characterized. A metazoan source of nuclei represents a direct and comprehensive way to identify and characterize critical tissue-specific variants of TFIID.

The ability to efficiently generate tissue-specific TFIID activities will also be a powerful tool for functional studies. With access to such material, it would be feasible to attempt recreation of tissue-specific transcription *in vitro*. Using similar complementation assays as described in section 4.3, holo-TFIID can be evaluated for regulatory functions by using promoters templates and/or activators that are known to be functionally restricted to certain tissues. For example, the ability of muscle-derived TFIID to activate transcription from a muscle-specific promoter could be contrasted to its ability to do such from a neuro-specific promoter. The results of such comparative functional studies could shed light on subtle mechanisms of tissue-specific transcription, mechanisms that have eluded past analysis using recombinant, highly purified factors and strong human viral promoters.

5.2.2 Other transcription systems

As mentioned earlier, one of the hallmarks of eukaryotic biology is the existence of three RNA polymerases. The majority of research efforts have focussed on mechanisms of protein-coding gene transcription, the unique obligation of RNA Pol II (see section 2.2). However, the other two polymerases, RNA Pol I and RNA Pol III have their own specialized functions and transcribe RNA from unique promoters. Like RNA Pol II, they require a battery of accessory factors analogous to the GTFs to be recruited accurately to their individual promoters. Although the functions of these three transcriptional mechanisms are mutually exclusive, the TBP factor is recognized to play a central role in

all three. Similar to the TFIID complex, TBP is a subunit in transcription factors that are vital to RNA polymerase recruitment for Pol I- and Pol III-specific promoters. For RNA Pol I promoters, the TBP protein combines with three other proteins (TAF_Is) to form a holo-complex known as SL-1 (148). An analogous complex for mammalian RNA Pol III is a complex known as TFIIIB, which is composed of TBP, a TFIIB-related factor called BRF1, and a partially defined activity referred to as Pol III-specific B''. A second RNA Pol III-specific complex, $SNAP_c$, is required for a subset of RNA Pol III promoters. Although distinct from TFIIIC, the $SNAP_c$ complex is also composed of TBP and an individual set of TAF_{III}s (149). Analogous to the TFIID complex, SL-1, TFIIIC, and $SNAP_c$ play pivotal roles in recruiting their respective polymerases to the correct promoters (26;61). Thus the transgenic model described in this thesis work could prove to be a useful approach for further examination of these other important transcriptional pathways. Tissue-derived transgenic hTBP should be evaluated for its ability to yield holo-complexes related to the regulated transcription from RNA Pol I and Pol III promoters. Such complexes would prove invaluable for qualitative and functional studies related to regulated expression of RNA Pol I and RNA Pol III genes.

5.2.3 Drug discovery

In addition to the contributions to basic transcription research that such a model offers, it is noteworthy that pharmaceutical applications have been anticipated for this transgenic mouse line as well. The myriad of mechanisms surrounding genetic regulation represent tempting possibilities for development of therapeutic agents. Such agents are envisioned to act as specific ligands, selectively antagonizing or enhancing targeted gene-specific molecular interactions. Historically, many synthetic compounds that have a recognized affinity for specific DNA sequences have been shown to selectively effect transcription factor function (150-152), including TBP-TATA interactions (153). The rapidly expanding body of knowledge of transcriptional activation mechanisms has more recently led some to consider gene-specific activator-coactivator interactions as possible targets of drug design (154). The tagged hTBP mouse model offers the potential characterization of novel, perhaps tissue-specific TAF moieties that can become targets for this approach.

More ambitiously would be the use of holo-complexes from this model to carry out screening for small molecules that selectively affect transcription. Parallel *in vitro* transcription assays could be carried out in the presence of various compounds, which would then be assessed for therapeutic value based on their relative abilities to antagonize (or enhance) regulated *in vitro* transcription. The use of transgenic, affinity-purified holo-TFIID adds the important elements of versatility and feasibility to such a screening approach. The search for agents could be biased toward specific tissue (and, thus, specific genes) by using holo-TFIID from selected tissue combined with tissue-specific promoter templates. Such an approach would be appropriate for a "high-throughput" primary screening strategy, where many compounds from a chemical library are tested in parallel. With sufficient technical support, it is conceivable that such *in vitro* transcription screening could be carried out automatically (robotically) to

test hundreds of thousands of compounds for possible candidate agents. Although the whole animal system described in this work may not be feasible to generate the amounts of holo-TFIID needed for such screening, it serves as a "proof of concept" that such a system is possible. The creation of a this transgenic system in a larger species (for example rat or swine) would likely prove more useful to support such screening.

6. CITED LITERATURE

1. Allison, L. A., M. Moyle, M. Shales, and C. J. Ingles. 1985. Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. *Cell* 42:599-610.

2. Young, R. A. 1991. RNA polymerase II. *Annual Review of Biochemistry* 60:689-715.

3. Woychik, N. A., S. M. Liao, P. A. Kolodziej, and R. A. Young. 1990. Subunits shared by eukaryotic nuclear RNA polymerases. *Genes and Development* 4:313-323.

4. McKune, K. and N. A. Woychik. 1994. Functional substitution of an essential yeast RNA polymerase subunit by a highly conserved mammalian counterpart. *Molecular and Cellular Biology* 14:4155-4159.

5. Weil, P. A., D. S. Luse, J. Segall, and R. G. Roeder. 1979. Selective and accurate initiation of transcription at the Ad2 major late promotor in a soluble system dependent on purified RNA polymerase II and DNA. *Cell* 18:469-484.

6. Helmann, J. D. Bacterial Sigma Factors, p.1-17. In *Transcription: Mechanisms and Regulation*, R. C. Conaway and J. W. Conaway (ed.), Raven Press (1994).

7. Buratowski, S., S. Hahn, L. Guarente, and P. A. Sharp. 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* 56:549-561.

8. Kaufmann, J., C. P. Verrijzer, J. Shao, and S. T. Smale. 1996. CIF, an essential cofactor for TFIID-dependent initiator function. *Genes and Development* 10:873-886.

9. Matsui, T., J. Segall, P. A. Weil, and R. G. Roeder. 1980. Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *Journal of Biological Chemistry* 255:11992-11996.

10. Kaiser, K. and M. Meisterernst. 1996. The human general co-factors. *Trends in Biochemical Sciences* 21:342-345.

11. Sawadogo, M. and R. G. Roeder. 1985. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell* 43:165-175.

12. Smale, S. T. 1997. Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. *Biochim. Biophys. Acta* 1351:73-88.

13. Roeder, R. G. 1996. The role of general initiation factors in transcription by RNA polymerase II. *Trends in Biochemical Sciences* 21:327-335.

14. Prioleau, M. N., J. Huet, A. Sentenac, and M. Mechali. 1994. Competition between chromatin and transcription complex assembly regulates gene expression during early development. *Cell* 77:439-449.

15. Chatterjee, S. and K. Struhl. 1995. Connecting a promoter-bound protein to TBP bypasses the need for a transcriptional activation domain. *Nature* 374:820-822.

16. Klages, N. and M. Strubin. 1995. Stimulation of RNA polymerase II transcription initiation by recruitment of TBP in vivo. *Nature* 374:822-823.

17. Xiao, H., J. D. Friesen, and J. T. Lis. 1995. Recruiting TATA-binding protein to a promoter: transcriptional activation without an upstream activator. *Molecular and Cellular Biology* 15:5757-5761.

18. Starr, D. B. and D. K. Hawley. 1991. TFIID binds in the minor groove of the TATA box. *Cell* 67:1231-1240.

19. Lee, D. K., M. Horikoshi, and R. G. Roeder. 1991. Interaction of TFIID in the minor groove of the TATA element. *Cell* 67:1241-1250.

20. Horikoshi, M., C. Bernuccioli, R. Takada, J. Wang, T. Yamamoto, and R. G. Roeder. 1992. Transcription factor TFIID induces DNA bending upon binding to the TATA element. *Proceedings of the National Academy of Sciences, U.S.A.* 89:1060-1064.

21. Jacobson, R. H. and R. Tjian. 1996. Transcription factor IIA: a structure with multiple functions. *Science* 272:827-828.

22. Nikolov, D. B. and S. K. Burley. 1997. RNA polymerase II transcription initiation: a structural view. *Proceedings of the National Academy of Sciences, U.S.A.* 94:15-22.

23. Imbalzano, A. N., K. S. Zaret, and R. E. Kingston. 1994. Transcription factor (TF) IIB and TFIIA can independently increase the affinity of the TATA-binding protein for DNA. *Journal of Biological Chemistry* 269:8280-8286.

24. Kim, T. K., S. Hashimoto, R. J. Kelleher, P. M. Flanagan, R. D. Kornberg, M. Horikoshi, and R. G. Roeder. 1994. Effects of activation-defective TBP mutations on transcription initiation in yeast. *Nature* 369:252-255.

25. Ha, I., S. Roberts, E. Maldonado, X. Sun, L. U. Kim, M. Green, and D. Reinberg. 1993. Multiple functional domains of human transcription factor IIB: distinct interactions with two general transcription factors and RNA polymerase II. *Genes and Development* 7:1021-1032.

26. Zawel, L. and D. Reinberg. 1995. Common themes in assembly and function of eukaryotic transcription complexes. *Annual Review of Biochemistry* 64:533-561.

27. Leuther, K. K., D. A. Bushnell, and R. D. Kornberg. 1996. Two-dimensional crystallography of TFIIB- and IIE-RNA polymerase II complexes: implications for start site selection and initiation complex formation. *Cell* 85:773-779.

28. Parvin, J. D. and P. A. Sharp. 1993. DNA topology and a minimal set of basal factors for transcription by RNA polymerase II. *Cell* 73:533-540.

29. Tan, S., R. C. Conaway, and J. W. Conaway. 1995. Dissection of transcription factor TFIIF functional domains required for initiation and elongation. *Proceedings of the National Academy of Sciences, U.S.A.* 92:6042-6046.

30. Killeen, M. T. and J. F. Greenblatt. 1992. The general transcription factor RAP30 binds to RNA polymerase II and prevents it from binding nonspecifically to DNA. *Molecular and Cellular Biology* 12:30-37.

31. Conaway, R. C. and J. W. Conaway. 1993. General initiation factors for RNA polymerase II. *Annual Review of Biochemistry* 62:161-190.

32. Zawel, L. and D. Reinberg. 1993. Initiation of transcription by RNA polymerase II: a multi-step process. *Prog. Nucleic. Acid. Res. Mol. Biol.* 44:67-108.

33. Ohkuma, Y. and R. G. Roeder. 1994. Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. *Nature* 368:160-163.

34. Maxon, M. E., J. A. Goodrich, and R. Tjian. 1994. Transcription factor IIE binds preferentially to RNA polymerase IIa and recruits TFIIH: a model for promoter clearance. *Genes and Development* 8:515-524.

35. Goodrich, J. A. and R. Tjian. 1994. Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell* 77:145-156.

36. Ohkuma, Y., S. Hashimoto, C. K. Wang, M. Horikoshi, and R. G. Roeder. 1995. Analysis of the role of TFIIE in basal transcription and TFIIH-mediated carboxy-terminal domain phosphorylation through structure-function studies of TFIIE-alpha. *Molecular and Cellular Biology* 15:4856-4866.

37. Drapkin, R. and D. Reinberg. 1994. The multifunctional TFIIH complex and transcriptional control. *Trends in Biochemical Sciences* 19:504-508.

38. Svejstrup, J. Q., P. Vichi, and J. M. Egly. 1996. The multiple roles of transcription/repair factor TFIIH. *Trends in Biochemical Sciences* 21:346-350.

39. Svejstrup, J. Q., Z. Wang, W. J. Feaver, X. Wu, D. A. Bushnell, T. F. Donahue, E. C. Friedberg, and R. D. Kornberg. 1995. Different forms of TFIIH for transcription and DNA repair: holo-TFIIH and a nucleotide excision repairosome. *Cell* 80:21-28.

40. Holstege, F. C., P. C. van der Vliet, and H. T. Timmers. 1996. Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. *EMBO Journal* 15:1666-1677.

41. Dahmus, M. E. 1996. Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *Journal of Biological Chemistry* 271:19009-19012.

42. Zawel, L., K. P. Kumar, and D. Reinberg. 1995. Recycling of the general transcription factors during RNA polymerase II transcription. *Genes and Development* 9:1479-1490.

43. Van-Dyke, M. W., R. G. Roeder, and M. Sawadogo. 1988. Physical analysis of transcription preinitiation complex assembly on a class II gene promoter. *Science* 241:1335-1338.

44. Van-Dyke, M. W., M. Sawadogo, and R. G. Roeder. 1989. Stability of transcription complexes on class II genes. *Molecular and Cellular Biology* 9:342-344.

45. Koleske, A. J. and R. A. Young. 1995. The RNA polymerase II holoenzyme and its implications for gene regulation. *Trends in Biochemical Sciences* 20:113-116.

46. Halle, J. P. and M. Meisterernst. 1996. Gene expression: increasing evidence for a transcriptosome. *Trends in Genetics* 12:161-163.

47. Gaudreau, L., M. Adam, and M. Ptashne. 1998. Activation of transcription in vitro by recruitment of the yeast RNA polymerase II holoenzyme. *Molecular Cell* 1:913-916.

48. Horikoshi, M., M. F. Carey, H. Kakidani, and R. G. Roeder. 1988. Mechanism of action of a yeast activator: direct effect of GAL4 derivatives on mammalian TFIID-promoter interactions. *Cell* 54:665-669.

49. Pugh, B. F. and R. Tjian. 1990. Mechanism of transcriptional activation by Sp1: evidence for coactivators. *Cell* 61:1187-1197.

50. Hahn, S., S. Buratowski, P. A. Sharp, and L. Guarente. 1989. Isolation of the gene encoding the yeast TATA binding protein TFIID: a gene identical to the SPT15 suppressor of Ty element insertions. *Cell* 58:1173-1181.

51. Cavallini, B., I. Faus, H. Matthes, J. M. Chipoulet, B. Winsor, J. M. Egly, and P. Chambon. 1989. Cloning of the gene encoding the yeast protein BTF1Y, which can substitute for the human TATA box-binding factor. *Proceedings of the National Academy of Sciences, U.S.A.* 86:9803-9807.

52. Schmidt, M. C., C. C. Kao, R. Pei, and A. J. Berk. 1989. Yeast TATA-box transcription factor gene. *Proceedings of the National Academy of Sciences, U.S.A.* 86:7785-7789.

53. Horikoshi, M., C. K. Wang, H. Fujii, J. A. Cromlish, P. A. Weil, and R. G. Roeder. 1989. Cloning and structure of a yeast gene encoding a general transcription initiation factor TFIID that binds to the TATA box. *Nature* 341:299-303.

54. Hoey, T., B. D. Dynlacht, M. G. Peterson, B. F. Pugh, and R. Tjian. 1990. Isolation and characterization of the Drosophila gene encoding the TATA box binding protein, TFIID. *Cell* 61:1179-1186.

55. Muhich, M. L., C. T. Iida, M. Horikoshi, R. G. Roeder, and C. S. Parker. 1990. cDNA clone encoding Drosophila transcription factor TFIID. *Proceedings of the National Academy of Sciences, U.S.A.* 87:9148-9152.

56. Hoffman, A., E. Sinn, T. Yamamoto, J. Wang, A. Roy, M. Horikoshi, and R. G. Roeder. 1990. Highly conserved core domain and unique N terminus with presumptive regulatory motifs in a human TATA factor (TFIID). *Nature* 346:387-390.

57. Kao, C. C., P. M. Lieberman, M. C. Schmidt, Q. Zhou, R. Pei, and A. J. Berk. 1990. Cloning of a transcriptionally active human TATA binding factor. *Science* 248:1646-1650.

58. Peterson, M. G., N. Tanese, B. F. Pugh, and R. Tjian. 1990. Functional domains and upstream activation properties of cloned human TATA binding protein. *Science* 248:1625-1630.

59. Tamura, T., K. Sumita, I. Fujino, A. Aoyama, M. Horikoshi, A. Hoffmann, R. G. Roeder, M. Muramatsu, and K. Mikoshiba. 1991. Striking homology of the 'variable' N-terminal as well as the 'conserved core' domains of the mouse and human TATA-factors (TFIID). *Nucleic Acids Research* 19:3861-3865.

60. Nikolov, D. B. and S. K. Burley. 1994. 2.1 Å resolution refined structure of a TATA box-binding protein (TBP). *Nat. Struct. Biol.* 1:621-637.

61. Hernandez, N. 1993. TBP, a universal eukaryotic transcription factor? *Genes and Development* 7:1291-1308.

62. Moncollin, V., R. D. Roy, and J.-M. Egly. The TATA Saga: Structure and Function of the TATA-Binding Factor, *Transcription: Mechanisms and Regulation*, R. C. a. C. J. W. Conaway (ed.), Raven Press (1994).

63. Kim, J. L., D. B. Nikolov, and S. K. Burley. 1993. Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* 365:520-527.

64. Kim, J. L. and S. K. Burley. 1994. 1.9 A resolution refined structure of TBP recognizing the minor groove of TATAAAAG. *Nat. Struct. Biol.* 1:638-653.

65. Kim, Y., J. H. Geiger, S. Hahn, and P. B. Sigler. 1993. Crystal structure of a yeast TBP/TATA-box complex. *Nature* 365:512-520.

66. Nikolov, D. B., H. Chen, E. D. Halay, A. Hoffman, R. G. Roeder, and S. K. Burley. 1996. Crystal structure of a human TATA box-binding protein/TATA element complex. *Proceedings of the National Academy of Sciences, U.S.A.* 93:4862-4867.

67. Dynlacht, B. D., T. Hoey, and R. Tjian. 1991. Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* 66:563-576.

68. Tanese, N., B. F. Pugh, and R. Tjian. 1991. Coactivators for a proline-rich activator purified from the multisubunit human TFIID complex. *Genes and Development* 5:2212-2224.

69. Reese, J. C., L. Apone, S. S. Walker, L. A. Griffin, and M. R. Green. 1994. Yeast $TAF_{II}S$ in a multisubunit complex required for activated transcription. *Nature* 371:523-527.

70. Poon, D., Y. Bai, A. M. Campbell, S. Bjorklund, Y. J. Kim, S. Zhou, R. D. Kornberg, and P. A. Weil. 1995. Identification and characterization of a TFIID-like multiprotein complex from Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences, U.S.A.* 92:8224-8228.

71. Zhou, Q., P. M. Lieberman, T. G. Boyer, and A. J. Berk. 1992. Holo-TFIID supports transcriptional stimulation by diverse activators and from a TATA-less promoter. *Genes and Development* 6:1964-1974.

72. Chiang, C. M., H. Ge, Z. Wang, A. Hoffmann, and R. G. Roeder. 1993. Unique TATA-binding protein-containing complexes and cofactors involved in transcription by RNA polymerases II and III. *EMBO Journal* 12:2749-2762.

73. Takada, R., Y. Nakatani, A. Hoffmann, T. Kokubo, S. Hasegawa, R. G. Roeder, and M. Horikoshi. 1992. Identification of human TFIID components and direct interaction between a 250-kDa polypeptide and the TATA box-binding protein (TFIID tau). *Proceedings of the National Academy of Sciences, U.S.A.* 89:11809-11813.

74. Kokubo, T., R. Takada, S. Yamashita, D. W. Gong, R. G. Roeder, M. Horikoshi, and Y. Nakatani. 1993. Identification of TFIID components required for transcriptional activation by upstream stimulatory factor. *Journal of Biological Chemistry* 268:17554-17558.

75. Brou, C., S. Chaudhary, I. Davidson, Y. Lutz, J. Wu, J. M. Egly, L. Tora, and P. Chambon. 1993. Distinct TFIID complexes mediate the effect of different transcriptional activators. *EMBO Journal* 12:489-499.

76. Zhou, Q. and A. J. Berk. 1995. The yeast TATA-binding protein (TBP) core domain assembles with human TBP-associated factors into a functional TFIID complex. *Molecular and Cellular Biology* 15:534-539.

77. Verrijzer, C. P., K. Yokomori, J. L. Chen, and R. Tjian. 1994. Drosophila $TAF_{II}150$: similarity to yeast gene TSM-1 and specific binding to core promoter DNA. *Science* 264:933-941.

78. Moqtaderi, Z., J. D. Yale, K. Struhl, and S. Buratowski. 1996. Yeast homologues of higher eukaryotic TFIID subunits. *Proceedings of the National Academy of Sciences, U.S.A.* 93:14654-14658.

79. Hoey, T., R. O. Weinzierl, G. Gill, J. L. Chen, B. D. Dynlacht, and R. Tjian. 1993. Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators. *Cell* 72:247-260.

80. Saluja, D., M. F. Vassallo, and N. Tanese. 1998. Distinct subdomains of human $TAF_{II}130$ are required for interactions with glutamine-rich transcriptional activators. *Molecular and Cellular Biology* 18:5734-5743.

81. Ferreri, K., G. Gill, and M. Montminy. 1994. The cAMP-regulated transcription factor CREB interacts with a component of the TFIID complex. *Proceedings of the National Academy of Sciences, U.S.A.* 91:1210-1213.

82. Farmer, G., J. Colgan, Y. Nakatani, J. L. Manley, and C. Prives. 1996. Functional interaction between p53, the TATA-binding protein (TBP), andTBP-associated factors in vivo. *Molecular and Cellular Biology* 16:4295-4304.

83. Thut, C. J., J. L. Chen, R. Klemm, and R. Tjian. 1995. p53 transcriptional activation mediated by coactivators $TAF_{II}40$ and $TAF_{II}60$. *Science* 267:100-104.

84. Wang, E. H., S. Zou, and R. Tjian. 1997. TAF(II)250-dependent transcription of cyclin A is directed by ATF activator proteins. *Genes and Development* 11:2658-2669.

85. Jacq, X., C. Brou, Y. Lutz, I. Davidson, P. Chambon, and L. Tora . 1994. Human $TAF_{II}30$ is present in a distinct TFIID complex and is required for transcriptional activation by the estrogen receptor. *Cell* 79:107-117.

86. May, M., G. Mengus, A. C. Lavigne, P. Chambon, and I. Davidson. 1996. Human TAF(II28) promotes transcriptional stimulation by activation function 2 of the retinoid X receptors. *EMBO Journal*.15:3093-3104.

87. Mengus, G., M. May, L. Carre, P. Chambon, and I. Davidson. 1997. Human TAF(II)135 potentiates transcriptional activation by the AF-2s of the retinoic acid, vitamin D3, and thyroid hormone receptors in mammalian cells. *Genes and Development* 11:1381-1395.

88. Zhou, J., J. Zwicker, P. Szymanski, M. Levine, and R. Tjian. 1998. TAF_{II} mutations disrupt Dorsal activation in the *Drosophila* embryo. *Proceedings of the National Academy of Sciences, U.S.A.* 95:13483-13488.

89. Sauer, F., S. K. Hansen, and R. Tjian. 1995. Multiple TAF_{II}s directing synergistic activation of transcription [comment]. *Science* 270:1783-1788.

90. Sauer, F., S. K. Hansen, and R. Tjian. 1995. DNA template and activator-coactivator requirements for transcriptional synergism by Drosophila bicoid. *Science* 270:1825-1828.

91. Martinez, E., C. M. Chiang, H. Ge, and R. G. Roeder. 1994. TATA-binding protein-associated factor(s) in TFIID function through the initiator to direct basal transcription from a TATA-less class II promoter. *EMBO Journal* 13:3115-3126.

92. Verrijzer, C. P., J. L. Chen, K. Yokomori, and R. Tjian. 1995. Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. *Cell* 81:1115-1125.

93. Hansen, S. K. and R. Tjian. 1995. TAFs and TFIIA mediate differential utilization of the tandem Adh promoters. *Cell* 82:565-575.

94. Burke, T. W. and J. T. Kadonaga. 1996. Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. *Genes and Development* 10:711-724.

95. Burke, T. W. and J. T. Kadonaga. 1997. The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by $TAF_{II}60$ of Drosophila. *Genes and Development* 11:3020-3031.

96. Mizzen, C. A., X. J. Yang, T. Kokubo, J. E. Brownell, A. J. Bannister, T. Owen Hughes, J. Workman, L. Wang, S. L. Berger, T. Kouzarides, Y. Nakatani, and C. D. Allis. 1996. The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell* 87:1261-1270.

97. Dikstein, R., S. Ruppert, and R. Tjian. 1996. $TAF_{II}250$ is a bipartite protein kinase that phosphorylates the base transcription factor RAP74. *Cell* 84:781-790.

98. O'Brien, T. and R. Tjian. 1998. Functional analysis of the human $TAF_{II}250$ N-terminal kinase domain. *Molecular Cell* 1:905-911.

99. Hoffmann, A., C. M. Chiang, T. Oelgeschlager, X. Xie, S. K. Burley, Y. Nakatani, and R. G. Roeder. 1996. A histone octamer-like structure within TFIID. *Nature* 380:356-359.

100. Burley, S. K. and R. G. Roeder. 1996. Biochemistry and structural biology of transcription factor IID (TFIID). *Annual Review of Biochemistry* 65:769-799.

101. Oelgeschlager, T., C. M. Chiang, and R. G. Roeder. 1996. Topology and reorganization of a human TFIID-promoter complex. *Nature* 382:735-738.

102. Hoffmann, A., T. Oelgeschlager, and R. G. Roeder. 1997. Considerations of transcriptional control mechanisms: do TFIID-core promoter complexes recapitulate nucleosome-like functions? *Proceedings of the National Academy of Sciences, U.S.A.* 94:8928-8935.

103. Verrijzer, C. P. and R. Tjian. 1996. TAFs mediate transcriptional activation and promoter selectivity. *Trends in Biochemical Sciences* 21:338-342.

104. Brou, C., J. Wu, S. Ali, E. Scheer, C. Lang, I. Davidson, P. Chambon, and L. Tora. 1993. Different TBP-associated factors are required for mediating the stimulation of transcription in vitro by the acidic transactivator GAL-VP16 and the two nonacidic activation functions of the estrogen receptor. *Nucleic Acids Research* 21:5-12.

105. Dubrovskaya, V., A. C. Lavigne, I. Davidson, J. Acker, A. Staub, and L. Tora. 1996. Distinct domains of $hTAF_{II}100$ are required for functional interaction with transcription factor TFIIF beta (RAP30) and incorporation into the TFIID complex. *EMBO Journal* 15:3702-3712.

106. Mengus, G., M. May, X. Jacq, A. Staub, L. Tora, P. Chambon, and I. Davidson. 1995. Cloning and characterization of $hTAF_{II}18$, $hTAF_{II}20$ and $hTAF_{II}28$: three subunits of the human transcription factor TFIID. *EMBO Journal* 14:1520-1531.

107. Attardi, L. D. and R. Tjian. 1993. Drosophila tissue-specific transcription factor NTF-1 contains a novel isoleucine-rich activation motif. *Genes and Development* 7:1341-1353.

108. Kunzler, M., G. H. Braus, O. Georgiev, K. Seipel, and W. Schaffner. 1994. Functional differences between mammalian transcription activation domains at the yeast GAL1 promoter. *EMBO Journal* 13:641-645.

109. Ponticelli, A. S., T. S. Pardee, and K. Struhl. 1995. The glutamine-rich activation domains of human Sp1 do not stimulate transcription in Saccharomyces cerevisiae. *Molecular and Cellular Biology* 15:983-988.

110. Ge, H., E. Martinez, C. M. Chiang, and R. G. Roeder. 1996. Activator-dependent transcription by mammalian RNA polymerase II: in vitro reconstitution with general transcription factors and cofactors. *Methods in Enzymology* 274:57-71.

111. Dikstein, R., S. Zhou, and R. Tjian. 1996. Human $TAF_{II}105$ is a cell type-specific TFIID subunit related to $hTAF_{II}130$. *Cell* 87:137-146.

112. Yamit, H. A. and R. Dikstein. 1998. $TAF_{II}105$ mediates activation of anti-apoptotic genes by NF-kappaB. *EMBO Journal* 17:5161-5169.

113. Purrello, M., C. Di-Pietro, A. Viola, A. Rapisarda, S. Stevens, M. Guermah, Y. Tao, C. Bonaiuto, A. Arcidiacono, A. Messina, G. Sichel, K. H. Grzeschik, and R. Roeder. 1998. Genomics and transcription analysis of human TFIID. *Oncogene* 16:1633-1638.

114. Gottesfeld, J. M. and D. J. Forbes. 1997. Mitotic repression of the transcriptional machinery. *Trends in Biochemical Sciences* 22:197-202.

115. Dynlacht, B. D. 1997. Regulation of transcription by proteins that control the cell cycle. *Nature* 389:149-152.

116. White, R. J., T. M. Gottlieb, C. S. Downes, and S. P. Jackson. 1995. Mitotic regulation of a TATA-binding-protein-containing complex. *Molecular and Cellular Biology* 15:1983-1992.

117. Segil, N., M. Guermah, A. Hoffmann, R. G. Roeder, and N. Heintz. 1996. Mitotic regulation of TFIID: inhibition of activator-dependent transcription and changes in subcellular localization. *Genes and Development* 10:2389-2400.

118. Conaway, R. C., D. Reines, K. P. Garrett, W. Powell, and J. W. Conaway. 1996. Purification of RNA polymerase II general transcription factors from rat liver. *Methods in Enzymology* 273:194-207.

119. Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. D. Seidman, J. A. Smith, and K. Struhl. *Current Protocols in Molecular Biology*, John Wiley & Sons (1995).

120. Deng, W. P. and J. A. Nickoloff. 1992. Site-directed mutagenesis of virtually any plasmid by eliminating a unique site. *Anal. Biochem.* 200:81-88.

121. Polites, H. G. and C. A. Pinkert. DNA Microinjection and Transgenic Animal Production, p.16-35. In *Transgenic Animal Technology: A Laboratory Handbook*, Carl A.Pinkert (ed.), Academic Press, Inc. (1994).

122. Kretzschmar, M., G. Stelzer, R. G. Roeder, and M. Meisterernst. 1994. RNA polymerase II cofactor PC2 facilitates activation of transcription by GAL4-AH in vitro. *Molecular and Cellular Biology* 14:3927-3937.

123. Meisterernst, M., A. L. Roy, H. M. Lieu, and R. G. Roeder. 1991. Activation of class II gene transcription by regulatory factors is potentiated by a novel activity. *Cell* 66:981-993.

124. Flores, O., H. Lu, and D. Reinberg. 1992. Factors involved in specific transcription by mammalian RNA polymerase II. Identification and characterization of factor IIH. *Journal of Biological Chemistry* 267:2786-2793.

125. Kim, W. Y. and M. E. Dahmus. 1988. Purification of RNA polymerase IIO from calf thymus. *Journal of Biological Chemistry* 263:18880-18885.

126. Werten, S., F. W. Langen, R. van-Schaik, H. T. Timmers, M. Meisterernst, and d. van, V. 1998. High-affinity DNA binding by the C-terminal domain of the transcriptional coactivator PC4 requires simultaneous interaction with two opposing unpaired strands and results in helix destabilization. *Journal of Molecular Biology* 276:367-377.

127. Werten, S., G. Stelzer, A. Goppelt, F. M. Langen, P. Gros, H. T. Timmers, d. van, V, and M. Meisterernst. 1998. Interaction of PC4 with melted DNA inhibits transcription. *EMBO Journal* 17:5103-5111.

128. Lin, Y. S., M. F. Carey, M. Ptashne, and M. R. Green. 1988. GAL4 derivatives function alone and synergistically with mammalian activators in vitro. *Cell* 54:659-664.

129. Sawadogo, M. and R. G. Roeder. 1985. Factors involved in specific transcription by human RNA polymerase II: analysis by a rapid and quantitative in vitro assay. *Proceedings of the National Academy of Sciences, U.S.A.* 82:4394-4398.

130. Wilson, I. A., H. L. Niman, R. A. Houghten, A. R. Cherenson, M. L. Connolly, and R. A. Lerner. 1984. The structure of an antigenic determinant in a protein. *Cell* 37:767-778.

131. Kolodziej, P. A. and R. A. Young. 1991. Epitope tagging and protein surveillance. *Methods in Enzymology* 194:508-519.

132. Zhou, Q., T. G. Boyer, and A. J. Berk. 1993. Factors (TAFs) required for activated transcription interact with TATA box-binding protein conserved core domain. *Genes and Development* 7:180-187.

133. Colgan, J. and J. L. Manley. 1992. TFIID can be rate limiting in vivo for TATA-containing, but not TATA-lacking, RNA polymerase II promoters. *Genes and Development* 6:304-315.

134. Hochuli, E., H. Dobeli, and A. Schacher. 1987. New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *Journal of Chromatography* 411:177-184.

135. Crowe, J., H. Dobeli, R. Gentz, E. Hochuli, D. Stuber, and K. Henco. 1994. 6xHis-Ni-NTA chromatography as a superior technique in recombinant protein expression/purification. *Methods Mol. Biol.* 31:371-387.

136. Porath, J. 1992. Immobilized metal ion affinity chromatography. *Protein Expression and Purification* 3:263-281.

137. Glanville, N., D. M. Durnam, and R. D. Palmiter. 1981. Structure of mouse metallothionein-I gene and its mRNA. *Nature* 292:267-269.

138. Searle, P. F., B. L. Davison, G. W. Stuart, T. M. Wilkie, G. Norstedt, and R. D. Palmiter. 1984. Regulation, linkage, and sequence of mouse metallothionein I and II genes. *Molecular and Cellular Biology* 4:1221-1230.

139. Ruther, U., C. Garber, D. Komitowski, R. Muller, and E. F. Wagner. 1987. Deregulated c-fos expression interferes with normal bone development in transgenic mice. *Nature* 325:412-416.

140. Erickson, L. A., G. J. Fici, J. E. Lund, T. P. Boyle, H. G. Polites, and K. R. Marotti. 1990. Development of venous occlusions in mice transgenic for the plasminogen activator inhibitor-1 gene. *Nature* 346:74-76.

141. Hofmann, S. L., D. W. Russell, M. S. Brown, J. L. Goldstein, and R. E. Hammer. 1988. Overexpression of low density lipoprotein (LDL) receptor eliminates LDL from plasma in transgenic mice. *Science* 239:1277-1281.

142. Palmiter, R. D., G. Norstedt, R. E. Gelinas, R. E. Hammer, and R. L. Brinster. 1983. Metallothionein-human GH fusion genes stimulate growth of mice. *Science* 222:809-814.

143. Castillo, M. B., M. R. Celio, C. Andressen, V. Gotzos, T. Rulicke, M. C. Berger, J. Weber, and M. W. Berchtold. 1995. Production and analysis of transgenic mice with ectopic expression of parvalbumin. *Arch. Biochem. Biophys.* 317:292-298.

144. Uetsuki, T., A. Naito, S. Nagata, and Y. Kaziro. 1989. Isolation and characterization of the human chromosomal gene for polypeptide chain elongation factor-1 alpha. *Journal of Biological Chemistry* 264:5791-5798.

145. Hanaoka, K., M. Hayasaka, T. Uetsuki, S. A. Fujisawa, and Y. Nabeshima. 1991. A stable cellular marker for the analysis of mouse chimeras: the bacterial chloramphenicol acetyltransferase gene driven by the human elongation factor 1 alpha promoter. *Differentiation.* 48:183-189.

146. Kim, D. W., T. Uetsuki, Y. Kaziro, N. Yamaguchi, and S. Sugano. 1990. Use of the human elongation factor 1 alpha promoter as a versatile and efficient expression system. *Gene* 91:217-223.

147. Mizushima, S. and S. Nagata. 1990. pEF-BOS, a powerful mammalian expression vector. *Nucleic Acids Research* 18:5322

148. Comai, L., N. Tanese, and R. Tjian. 1992. The TATA-binding protein and associated factors are integral components of the RNA polymerase I transcription factor, SL1. *Cell* 68:965-976.

149. Goodrich, J. A. and R. Tjian. 1994. TBP-TAF complexes: selectivity factors for eukaryotic transcription. *Current Opinion in Cell Biology* 6:403-409.

150. Ho, S. N., S. H. Boyer, S. L. Schreiber, S. J. Danishefsky, and G. R. Crabtree. 1994. Specific inhibition of formation of transcription complexes by a calicheamicin oligosaccharide: a paradigm for the development of transcriptional antagonists. *Proceedings of the National Academy of Sciences, U.S.A.* 91:9203-9207.

151. Liu, C., B. M. Smith, K. Ajito, H. Komatsu, P. L. Gomez, T. Li, E. A. Theodorakis, K. C. Nicolaou, and P. K. Vogt. 1996. Sequence-selective carbohydrate-DNA interaction: dimeric and monomeric forms of the calicheamicin oligosaccharide interfere with transcription factor function. *Proceedings of the National Academy of Sciences, U.S.A.* 93:940-944.

152. Gottesfeld, J. M., L. Neely, J. W. Trauger, E. E. Baird, and P. B. Dervan. 1997. Regulation of gene expression by small molecules. *Nature* 387:202-205.

153. Sun, D. and L. H. Hurley. 1995. TBP binding to the TATA box induces a specific downstream unwinding site that is targeted by pluramycin. *Chem. Biol.* 2:457-469.

154. Nyanguile, O., M. Uesugi, D. J. Austin, and G. L. Verdine. 1997. A nonnatural transcriptional coactivator. *Proceedings of the National Academy of Sciences, U.S.A.* 94:13402-13406.

7. Abbreviations

AdML	Adenovirus middle late
ATP	Adenosine 5'-triphosphate
BCD	Bicoid
bp	Base pairs
BSA	Bovine serum albumin
BV	Bed volume
cDNA	Copy DNA
CREB	cAMP responsive element binding protein
CTD	Carboxy terminal domain
CTP	Cytidine 5'-triphosphate
dATP	2'-Deoxyadenosine 5'-triphosphate
dCTP	2'-Deoxycytidine 5'-triphosphate
DEPC	Diethyl pyrocarbonate
dGTP	2'-Deoxyguanosine 5'-triphosphate
DNA	Deoxyribonucleic acid
DPE	Downstream promoter element
DTT	Dithiothreitol
dTTP	2'-Deoxythymidine 5'-triphosphate
EDTA	Ethylenediaminetetraacetic acid
EF	Elongation factor
EGTA	Ethylene glycol-bis[beta-aminoethyl ether]-N,N,N',N'-tetraacetic acid
G_0	F_0 generation of transgenic animal line
G_1	F_1 generation of transgenic animal line
GTF	General transcription factor
GTP	Guanosine 5'-triphosphate
HA	Haemaaglutinin
HB	Hunchback
Hepes	N-[2-hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid]
IMAC	Immobilized metal affinity chromatography
Inr	Initiator
kDa	Kilodaltons
mAb	Monoclonal antibody
mRNA	Messenger RNA
MT	Metallothiomine
OD	Optical density

OmG	3'-O-methylguanosine-5'-triphosphate
^{32}P	Isotope Phosphorus-32
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PIC	Pre-initiation complex
Pipes	1,4-Piperazinediethanesulfonic acid
PMSF	Phenylmethylsulfonylfluoride
Pol	Polymerase
PVDF	Polyvinylidene difluoride
RNA	Ribonucleic acid
RNAse	Ribonuclease
rpm	Rotations per minute
rRNA	Ribosomal RNA
RXR	Retinoic acid receptor
SDS	Sodium dodecylsulfate
TAE	Tris-acetate-EDTA
TAF	TBP-associating factor
TBE	Tris-borate-EDTA
TBP	TATA-binding protein
TE	Tris-EDTA
TEMED	N,N, N',N'-Tetramethylethylenediamine
TF	Transcription factor
Tris	Trihydroxymethylaminomethane
tRNA	Transfer RNA
UTP	Uridine 5'-triphosphate
UV	Ultraviolet
v/v	Volume per volume
w/v	Weight per volume